

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
1	2754	100.0	2754	10	AE027170		

Description

AR068839	Sequence
AR070139	Sequence

M68971	Rat	hexokin
M68972	Rat	hexokin

AC098769 Rattus no

AJ238540 Mus muscu
Y11668 Mus muscu

AC116811 Mus muscu

Y11667 Mus musculus

AY082375 Rattus no
 I119605 Rattus no

X96699 M.musculus

NOTED FOR INFO

ALIGNMENTS

RESULT 1				
AF027179				
LOCUS	AF027179	2754 bp	mRNA	linear
DEFINITION	Rattus norvegicus mutant type II hexokinase mRNA, complete cds.			ROD 16-DEC-1997
ACCESSION	AF027179			
VERSION	AF027179.1	GI:2689657		
KEYWORDS				
SOURCE				
ORGANISM				
	Rattus norvegicus.			
	Rattus norvegicus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
	Rattus.			
REFERENCE	1 (bases 1 to 2754)			
AUTHORS	Mathupala, S.P.			
TITLE	Type II Hexokinase of Hepatoma AS-30D			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 2754)			
AUTHORS	Mathupala, S.P.			
TITLE	Direct Submission			
JOURNAL	Submitted (27-SEP-1997) Biological Chemistry, Johns Hopkins			
	University School of Medicine, 725 N. Wolfe Street, Baltimore, MD			
	21205, USA			
FEATURES				
source	location/Qualifiers			
	1..2754			
	/organism="Rattus norvegicus"			
	/strain="Sprague-Dawley"			
	/db_xref="taxon:10116"			
	/cell_line="AS-30D"			
	/cell_type="ascites hepatoma"			
	1..2754			
	/note="contains 4 amino acid mutations when compared to			
	skeletal muscle type II hexokinase"			
	/codon_start=1			
	/product="mutant type II hexokinase"			
	/protein_id="AB91396.1"			
	/db_xref="GI:2689658"			
	/translation="MASHIMICLTPELLNONQVKVDFLYHMRISPELTLSRRF			
	KREEMKIGATPTPAVKKMLPTFVSGSPDGRHEPEALDLSGNNRYLVLRVTDGGL			
	QRENMNLYALIEDIMRSGTQIFDHIAETLAKEMKQIT			

ETKDVSDIEEDKDGIEKAYOILMLRGLNPLEDCVATHRICQIVSTRASLCATLAA
VLMRLKENKGERLRKSTIGVDSQSYTKRPHKRAKVRRLVDDCVRFLESHEDSG
KCAAAVTVAAVLADQHRAROKTLESLSHOLEYVRKRVKMEQGLSKETHPV
VKMLPTVYCAPDGTKEGDFLADLGGNFRVLLRVNNGRRVGENKIKYSPOVY
MGTGBELFDHIVOC IADFLIEMGMKGVSLPGFSPCCQNSIDOS IILKWKGEK
ASGCGEDVYLLKLEAIHRRREFDLVAVAVNDVGTVMTCGYEDPHCEVLIVGTGS
NACVMEBNNEVLVDGEGRMCVNMEMGAFGNGCLDDLRVFDYAVDELILNPKOR
FKRAISGNILGICIVNIIILIDFTKRLGLFRGRISERLKTRGISELQISDCLAL
QVRAILRLGLSESTDDSIIVKEXCTVAVARRAOLCGAGMAVVDKIRENGLDNPKV
TVGVDDTLKLPHPAKVMEHETVRDLAKCDVSPFLESDSGSKAALLTAVACRIRFA
GOR"

variation
344
/note="nucleotide change resulting in an amino acid change
of leu instead of pro"
/replace="c"
575
/note="nucleotide change resulting in an amino acid change
of val instead of ala"
/replace="c"
2345
/note="nucleotide change resulting in an amino acid change
of ser instead of phe"
/replace="t"
2561
/note="nucleotide change resulting in an amino acid change
of pro instead of leu"
/replace="a"
BASE COUNT 642 a 659 c 869 g 584 t
ORIGIN

Query Match 100.0%; Score 2754; DB 10; Length 2754;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2754; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATCGCCCTGCATATGATGCCCTGCTTATTACGAGCTCAACCAACCAAGTGAG 60
DB 1 ATGATCGCCCTGCATATGATGCCCTGCTTATTACGAGCTCAACCAACCAAGTGAG 60
QY 61 AAGGTTGACCAATTTCTCTACCAACATGGGCTCTCAGATGAGACCCCTCTGAGATTCT 120
DB 61 AAGGTTGACCAATTTCTCTACCAACATGGGCTCTCAGATGAGACCCCTCTGAGATTCT 120
QY 121 AAGCGGTCGGAAGAGATGAGAAAGGCTAGAGACTACCAACCCACCTTACAGACT 180
DB 121 AAGCGGTCGGAAGAGATGAGAAAGGCTAGAGACTACCAACCCACCTTACAGACT 180
QY 181 GTGAATAATGTTGCTACCTTTGTGAGAGTCAACCTCGGATGGACAGAACATGGGAGTTCT 240
DB 181 GTGAATAATGTTGCTACCTTTGTGAGAGTCAACCTCGGATGGACAGAACATGGGAGTTCT 240
QY 241 CTGGCTCTGATCTTGGAGAACCACTTCGCTGCTCGAGTAAGGTGACGAGCAAT 300
DB 241 CTGGCTCTGATCTTGGAGAACCACTTCGCTGCTCGAGTAAGGTGACGAGCAAT 300
QY 301 GGCCCTCCAGAGAGTGGAGAACCAACATCTACGCCATCTTGAAGGACATCATGCGG 360
DB 301 GGCCCTCCAGAGAGTGGAGAACCAACATCTACGCCATCTTGAAGGACATCATGCGG 360
QY 361 GGCCTCCAGAGAGTGGAGAACCAACATCTACGCCATCTTGAAGGACATCATGCGG 420
DB 361 GGCCTCCAGAGAGTGGAGAACCAACATCTACGCCATCTTGAAGGACATCATGCGG 420
QY 421 CTACAAATCAAAAGAGAGAGTCCCTCTGAGTTTCACTTCTGCTTCCCTCCGACAG 480
DB 421 CTACAAATCAAAAGAGAGAGTCCCTCTGAGTTTCACTTCTGCTTCCCTCCGACAG 480
QY 481 ACAAAACGTGATGAGAGTTTGTCTGCTGACTAAGGGGTTCAAGTCAAGTGCGGT 540
DB 481 ACAAAACGTGATGAGAGTTTGTCTGCTGACTAAGGGGTTCAAGTCAAGTGCGGT 540
QY 541 GAAGGACAGAGATGTTGCTGACCTATCCGAGAGTTATCCAGCCCAAGGGGATTTGAC 600
DB 541 GAAGGACAGAGATGTTGCTGACCTATCCGAGAGTTATCCAGCCCAAGGGGATTTGAC 600

601 ATTGACATTTGTCGCCGCTGCTGATGATGACACAGTTGGGNCATGATCTTGCTATGAT 660
DB 601 ATTGACATTTGTCGCCGCTGCTGATGATGACACAGTTGGGNCATGATCTTGCTATGAT 660
QY 661 GATGACAGACTCGAGATTGGTCTCATTTGTGGGCACTGGACCAACGCCCTCTACATGAG 720
DB 661 GATGACAGACTCGAGATTGGTCTCATTTGTGGGCACTGGACCAACGCCCTCTACATGAG 720
QY 721 GAAATGCGTCATATTGATGATGCTGAGGAGAGATAGAGGGGCAATGTCATCAATGAG 780
DB 721 GAAATGCGTCATATTGATGATGCTGAGGAGAGATAGAGGGGCAATGTCATCAATGAG 780
QY 781 TGGGAGGCTTTGGGAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 TGGGAGGCTTTGGGAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 ATGACATGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 841 ATGACATGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 TACATGAGGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 TACATGAGGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 CAAGGAAACCTGACCCAGACTCTTACCACTGGCTCTTGAAGCAACCAAGATGATGAT 1020
DB 961 CAAGGAAACCTGACCCAGACTCTTACCACTGGCTCTTGAAGCAACCAAGATGATGAT 1020
QY 1021 GATATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 GATATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 CTGAATTCATTTGAGAGAGATTTGTGCGCAGCAGCAACCCAGATGATGATGATGAT 1140
DB 1081 CTGAATTCATTTGAGAGAGATTTGTGCGCAGCAGCAGCAACCCAGATGATGATGAT 1140
QY 1141 CGCTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 CGCTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 AAGGCGAGAGAGCTTGGCTCCACCATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 AAGGCGAGAGAGCTTGGCTCCACCATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 CCCCATTTTGCAGAGAGCTTGCATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 CCCCATTTTGCAGAGAGCTTGCATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 CGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1321 CGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 TACGCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 TACGCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1441 CACGAGAGCTTCTGAGAGTTAAGAGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGA 1500
DB 1441 CACGAGAGCTTCTGAGAGTTAAGAGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGA 1500
QY 1501 AAGGAG 1560
DB 1501 AAGGAG 1560
QY 1561 GATGAG 1620
DB 1561 GATGAG 1620
QY 1621 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
DB 1621 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680

```

QY 1681 TCATCCACAGAGGTTATGACATGGGAGAGGCTTTCAGACCATTTGTCAG 1740
Db 1681 TCATCCACAGAGGTTATGACATGGGAGAGGCTTTCAGACCATTTGTCAG 1740
QY 1741 TGCATTTGGGAGGCTTCTGAGTACATGGGAGGAGGCTGTCCTTTGGGGTTTC 1800
Db 1741 TGCATTTGGGAGGCTTCTGAGTACATGGGAGGAGGCTGTCCTTTGGGGTTTC 1800
QY 1801 ACATTTCTCTTCCCTTGGCAGACAGACGCTTACACAGACATCTCTCAAGTGACA 1860
Db 1801 ACATTTCTCTTCCCTTGGCAGACAGACGCTTACACAGACATCTCTCAAGTGACA 1860
QY 1861 AAGGATTTCAAGGATCTGCTGCTGAGGAGGATGAGTGTGCTCACTTGTGGAAGAGG 1920
Db 1861 AAGGATTTCAAGGATCTGCTGCTGAGGAGGAGGATGAGTGTGCTCACTTGTGGAAGAGG 1920
QY 1921 ATTACCGGCGAGAGAGGAGTGTGACCTGAGTGTGTTGCTGCTGCTGATACACAGTTGG 1980
Db 1921 ATTACCGGCGAGAGAGGAGTGTGACCTGAGTGTGTTGCTGCTGCTGATACACAGTTGG 1980
QY 1981 ACTATGATGACTTGTGCTTACAGAGACCTCACTGAGTGTGCTGCTGCTGCTGCTGCTG 2040
Db 1981 ACTATGATGACTTGTGCTTACAGAGACCTCACTGAGTGTGCTGCTGCTGCTGCTGCTG 2040
QY 2041 GGAAGCAAGGCTGCTACATGGAAGAGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 2041 GGAAGCAAGGCTGCTACATGGAAGAGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 2101 GGAGGATGCTGTCAACATGGAAGTGGGAGCATTTGGGAGCAATTTGCTGCTGCTGCTG 2160
Db 2101 GGAGGATGCTGTCAACATGGAAGTGGGAGCATTTGGGAGCAATTTGCTGCTGCTGCTG 2160
QY 2161 TTGCGGACCGTGTGTGATGTTGCTGTGATGAGTGTGCTGTCTCAACCTTGGCAAGAG 2220
Db 2161 TTGCGGACCGTGTGTGATGTTGCTGTGATGAGTGTGCTGTCTCAACCTTGGCAAGAG 2220
QY 2221 TTGCGAGATGATACAGGCGATCTACTTGGGAGATGTTGCGCAACATTTGCTGCTGCTG 2280
Db 2221 TTGCGAGATGATACAGGCGATCTACTTGGGAGATGTTGCGCAACATTTGCTGCTGCTG 2280
QY 2281 TTCAAGAGCGGGGCTGCTCTTCCGAGGCGGATCTCAAGAGGCTTCAAGAGAGGGA 2340
Db 2281 TTCAAGAGCGGGGCTGCTCTTCCGAGGCGGATCTCAAGAGGCTTCAAGAGAGGGA 2340
QY 2341 ATCTGAAACTAAGTCTCTCTCTCAAGATGAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTG 2400
Db 2341 ATCTGAAACTAAGTCTCTCTCTCAAGATGAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTG 2400
QY 2401 CGTGCATCTGCGGACCTAGAGGCTGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Db 2401 CGTGCATCTGCGGACCTAGAGGCTGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
QY 2461 GAGGTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
Db 2461 GAGGTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
QY 2521 GTAGTGTGACAGATAGAGAGAACGCTGCGGAGCAACCCCAAGTGTGAGTGTGAGGCTG 2580
Db 2521 GTAGTGTGACAGATAGAGAGAACGCTGCGGAGCAACCCCAAGTGTGAGTGTGAGGCTG 2580
QY 2581 GACGGGACTCTGTATAGCTTCTCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2640
Db 2581 GACGGGACTCTGTATAGCTTCTCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2640
QY 2641 GATGTGTGCTGCAAAATGTGAGTGTGCTTCTGGAATCCAGAGAGGAGGAGGAGGAGGAG 2700
Db 2641 GATGTGTGCTGCAAAATGTGAGTGTGCTTCTGGAATCCAGAGAGGAGGAGGAGGAGGAG 2700
QY 2701 GCAGTCTCATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2754
Db 2701 GCAGTCTCATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2754

```

```

RESULT 2
LOCUS AR068839 3635 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 15 from patent US 5854067.
ACCESSION AR068839
VERSION AR068839.1 GI:6001046
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3635)
AUTHORS Newgard, C. B., Han, H.-P. and Normington, K. D.
TITLE Hexokinase inhibitors
JOURNAL Patent: US 5854067-A 15 29-DEC-1998;
FEATURES
Source location/Qualifiers
1..3635
BASE COUNT 816 a 875 c 1098 g 846 t
ORIGIN
Query Match 92.6%; Score 2550; DB 6; Length 3635;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2750; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGATGCGCTCGCATATGATGCGCTGCTTATTCAGAGGCTCAACCAAAACAGTGACG 60
Db 198 ATGATGCGCTCGCATATGATGCGCTGCTTATTCAGAGGCTCAACCAAAACAGTGACG 257
QY 61 AAGGTTGACCAATTTCTTACCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 257
Db 258 AAGGTTGACCAATTTCTTACCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 317
QY 121 AAGCGTTCCGGAAGAGATGAGAAAGGCTGAGAGCTACACAGACCTTACAGAGCT 180
Db 318 AAGCGTTCCGGAAGAGATGAGAAAGGCTGAGAGCTACACAGACCTTACAGAGCT 377
QY 181 GTGAAATGTTGCTTACCTTGTGAGTCAACTCCGATGAGAGACAGATGAGAGGAGTTC 240
Db 378 GTGAAATGTTGCTTACCTTGTGAGTCAACTCCGATGAGAGACAGATGAGAGGAGTTC 437
QY 241 CTGCTGTGATCTTGGAGAACCACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 438 CTGCTGTGATCTTGGAGAACCACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 497
QY 301 GGCCTCCAGAGATGAGAGATGAGAACAGATCTACGCGATCTTGAAGAGATCATGCGG 360
Db 498 GGCCTCCAGAGATGAGAGATGAGAACAGATCTACGCGATCTTGAAGAGATCATGCGG 557
QY 361 GGCAGTGAACCAAGCTGTTGACCAACATGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTG 420
Db 558 GGCAGTGAACCAAGCTGTTGACCAACATGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTG 517
QY 421 CTGCAATCAAAAGAAAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 618 CTGCAATCAAAAGAAAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
QY 481 ACAAACCTGATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 678 ACAAACCTGATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
QY 541 GAAGCAGAGATGAGTGTGAGAGCTGATCCGAAAGTTATCCAGGCGAGAGGAGGAGGAGG 600
Db 738 GAAGCAGAGATGAGTGTGAGAGCTGATCCGAAAGTTATCCAGGCGAGAGGAGGAGGAGG 797
QY 601 ATTGACATTTGAGGCTGCTGATGACACAGTGTGAGACCATGATGATGATGATGATGAT 660
Db 798 ATTGACATTTGAGGCTGCTGATGACACAGTGTGAGACCATGATGATGATGATGATGAT 857
QY 661 GATCAGAACTGCGAGATTTGCTCATTTGTGGCACTGCGAGCAAGGCTGCTGCTGCTGCTG 720
Db 858 GATCAGAACTGCGAGATTTGCTCATTTGTGGCACTGCGAGCAAGGCTGCTGCTGCTGCTG 917
QY 721 GAAATGCTCATATTGACATGAGTGTGAGAGGAGATGAGAGGAGGAGGAGGAGGAGGAG 780

```

|||||
Db 918 GAAATGCGTATATGACATGTGGAGGAGATGAGGGCCATGTGATCAACATGGAG 977
781 TGGGAGCCCTTTGGGAGCAGCGTACATCATGTACATCCGAACGAGTTTGACGAGAG 840
978 TGGGAGCCCTTTGGGAGCAGCGTACATCATGTACATCCGAACGAGTTTGACGAGAG 1037
841 ATGACATGGGCTCGCTGAACCTTGGAAGCAGCTGTTTGAGACATGATTAGCGGAGTG 900
1038 ATGACATGGGCTCGCTGAACCTTGGAAGCAGCTGTTTGAGACATGATTAGCGGAGTG 1097
901 TACATGGGAGGCTGCTGATGCTCATCTGTGTAAGATGAGGCCAAGGACAGCTGTGTTT 960
1098 TACATGGGAGGCTGCTGATGCTCATCTGTGTAAGATGAGGCCAAGGACAGCTGTGTTT 1157
961 CAAGGGAACCTCAGGCCCAACTCTTACCACCTGCTCTTGAGACCAAAAGTGTCTG 1217
1158 CAAGGGAACCTCAGGCCCAACTCTTACCACCTGCTCTTGAGACCAAAAGTGTCTG 1217
1021 GATATTGAAGAGATTAAGATGGAATGGAAGGCTTACCAATCTGATGCGCTGGGT 1080
1218 GATATTGAAGAGATTAAGATGGAATGGAAGGCTTACCAATCTGATGCGCTGGGT 1277
1081 CTGATCCATTGCAAGAGATTTGTGGCCACGACCGAATCTGCGAGATTGTGTCCAG 1140
1278 CTGATCCATTGCAAGAGATTTGTGGCCACGACCGAATCTGCGAGATTGTGTCCAG 1337
1141 CGCTGGCCAGTCTGTGGGAGCCACCTTGCCGCGGTGCTGTGGCAATCAAGAGAAC 1200
1338 CGCTGGCCAGTCTGTGGGAGCCACCTTGCCGCGGTGCTGTGGCAATCAAGAGAAC 1397
1201 AAGGCGGAGAGCAGCTTGTGCTCACCATCGTGTGATGAGTGTGCTTACAAAGACAT 1260
1398 AAGGCGGAGAGCAGCTTGTGCTCACCATCGTGTGATGAGTGTGCTTACAAAGACAT 1457
1261 CCCCATTTTGGCAAGGCTCCATTAAGGACATGAGAGGCTGTGCCGACTGTGATCTC 1320
1458 CCCCATTTTGGCAAGGCTCCATTAAGGACATGAGAGGCTGTGCCGACTGTGATCTC 1517
1321 CGCTTCCTCCCTCTGAGATGAGGACGCGCAAGGGGCGCTATGTTGACGCGTGGCT 1380
1518 CGCTTCCTCCCTCTGAGATGAGGACGCGCAAGGGGCGCTATGTTGACGCGTGGCT 1577
1381 TACCGTCTGGCTGACCAACACCGGGCCGCAAGAACCTTGAGTCTGTGAAGCTGAGC 1440
1578 TACCGTCTGGCTGACCAACACCGGGCCGCAAGAACCTTGAGTCTGTGAAGCTGAGC 1637
1441 CACGAGCAGCTTGTGAGGTTTAAGAGAAATGAAGTGAATGAGCAGGGTCTGAGC 1500
1638 CACGAGCAGCTTGTGAGGTTTAAGAGAAATGAAGTGAATGAGCAGGGTCTGAGC 1697
1501 AAGGAGACGCTTGGGCTGCGCCCTGTGAAGATGCTGCCACTTACGTTGTGCCACTCCA 1560
1698 AAGGAGACGCTTGGGCTGCGCCCTGTGAAGATGCTGCCACTTACGTTGTGCCACTCCA 1757
1561 GATGGCAGAGAAAGAGACTTCTTGCCCTTGATCTTGGAGGAACAACTTCCGGGCTC 1620
1758 GATGGCAGAGAAAGAGACTTCTTGCCCTTGATCTTGGAGGAACAACTTCCGGGCTC 1817
1621 CTGCTGTGCTGCTGTAATGGAAGGAGGCGGTGAGATGCATTAACAGATCTAC 1680
1818 CTGCTGTGCTGCTGTAATGGAAGGAGGCGGTGAGATGCATTAACAGATCTAC 1877
1681 TCCATCCACAGAGAGATTATGATGACCTGGGGAAGGCTCTTGACACATTTGTCCAG 1740
1878 TCCATCCACAGAGAGATTATGATGACCTGGGGAAGGCTCTTGACACATTTGTCCAG 1937
1741 TGCATTGCGGAGCTTCTGAGATGACATGGGATGAAGGGGCTGTCCCTTGGGTTTC 1800
1938 TGCATTGCGGAGCTTCTGAGATGACATGGGATGAAGGGGCTGTCCCTTGGGTTTC 1997
1801 ACATTCTCCTTCCCTTGGCAGAGAACGCTTACAGCAGAGATCTCTCTCAAGTGAGCA 1860
|||||

Db 1998 ACATTCTCCTTCCCTTGGCAGAGAACGCTTACAGCAGAGATCTCTCTCAAGTGAGCA 2057
1861 AAGGATTCAGGATCTGCTGCTGCGAGGCTGAGATGTGTGTACCTTGTGAGGAAGCG 1920
2058 AAGGATTCAGGATCTGCTGCTGCGAGGCTGAGATGTGTGTACCTTGTGAGGAAGCG 2117
1921 ATTTCACCGGAGAGAGATTGACCTGATGTGTGTGCTGCTGCTGATGACAGACTTGGG 1980
2118 ATTTCACCGGAGAGAGATTGACCTGATGTGTGTGCTGCTGCTGATGACAGACTTGGG 2177
1981 ACTATGATGACTTGTGCTACGAAAGACCTCCTACTGTGAAGTGTGGCTCATTTGTCACC 2040
2178 ACTATGATGACTTGTGCTACGAAAGACCTCCTACTGTGAAGTGTGGCTCATTTGTCACC 2237
2041 GGAGCAACGCTGCTCATGAGAAAGATGCTTAATGTGAGTGTGTGAGCAGAGAGAG 2100
2238 GGAGCAACGCTGCTCATGAGAAAGATGCTTAATGTGAGTGTGTGAGCAGAGAGAG 2297
2101 GGAGGATGTGTCAACATGAGAGTGGGAGCATTTTGGGCAAAATGGCTGTGATGAC 2160
2298 GGAGGATGTGTCAACATGAGAGTGGGAGCATTTTGGGCAAAATGGCTGTGATGAC 2357
2161 TTGGCGACCGTGTGATGATGTTGCTGTGATGAGCTTCTCTCAACCTTGCAAAAGAGG 2220
2358 TTGGCGACCGTGTGATGATGTTGCTGTGATGAGCTTCTCTCAACCTTGCAAAAGAGG 2417
2221 TTGAGAAAGATGATCAGCGCATGCTTGTGGAGAGATTGTGGCCACATTTCTATCATGAT 2280
2418 TTGAGAAAGATGATCAGCGCATGCTTGTGGAGAGATTGTGGCCACATTTCTATCATGAT 2477
2281 TTGAGAAAGATGATCAGCGCATGCTTGTGGAGAGATTGTGGCCACATTTCTATCATGAT 2340
2478 TTGAGAAAGATGATCAGCGCATGCTTGTGGAGAGATTGTGGCCACATTTCTATCATGAT 2537
2341 ATCTTGAACATTAAGTCTGCTGCTGATGAGAGCAGCTGCTTCTCAAGCTTCAAGGTT 2400
2538 ATCTTGAACATTAAGTCTGCTGCTGATGAGAGCAGCTGCTTCTCAAGCTTCAAGGTT 2597
2401 CGTGCCATCTGCGCCACCTAAGGCTGTGAGAGCAGCTGTGATGAGAGATCATCGTGAAG 2460
2598 CGTGCCATCTGCGCCACCTAAGGCTGTGAGAGCAGCTGTGATGAGAGATCATCGTGAAG 2657
2461 GAGGTGTGCACTGTGTGTGCGCGCGGCTGACAGCTGTGTGGGCGACGATGCGCGCC 2520
2658 GAGGTGTGCACTGTGTGTGCGCGCGGCTGACAGCTGTGTGGGCGACGATGCGCGCC 2717
2521 GTAGTGACAAGATTAAGAGAAACCGTGGCTGTGACAAACCCCAAAATGACAGTGGCGTG 2580
2718 GTAGTGACAAGATTAAGAGAAACCGTGGCTGTGACAAACCCCAAAATGACAGTGGCGTG 2777
2581 GACGGGACTCTGTATTAAGCTTCACTTCCCTGCTGCAAGTGTGATGATGAGAGCGTGA 2640
2778 GACGGGACTCTGTATTAAGCTTCACTTCCCTGCTGCAAGTGTGATGATGAGAGCGTGA 2837
2641 GATCGGCTCGAAATGTGACGTGTCTTCTGTAATCCGAGGACGCTATGGGAAGGGA 2700
2838 GATCGGCTCGAAATGTGACGTGTCTTCTGTAATCCGAGGACGCTATGGGAAGGGA 2897
2701 GCAGCTCTCATACCTGCGGCTGCGGCAATCCGGAAGGCTGGGCAAGATAG 2754
2898 GCAGCTCTCATACCTGCGGCTGCGGCAATCCGGAAGGCTGGGCAAGATAG 2951

RESULT 3
LOCUS AR070139 3635 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 15 from patent US 5891717.
ACCESSION AR070139
VERSION AR070139.1 GI:7221027
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 3635)
AUTHORS Newgard,C.B., Han,H.-P., Becker,T.C. and Willson,J.E.
TITLE Methods and compositions for inhibiting hexokinase
JOURNAL Patent: US 5891717-A 15 06-Apr-1999;
FEATURES Location/Qualifiers
Source 1. 3635
BASE COUNT 816 a /organism="unknown"
ORIGIN 875 c 1098 g 846 t

Query Match 92.6%; Score 2550; DB 6; Length 3635;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2750; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGATCGCGCTGCGATATGCGCTGCTTATTCACGSGAGCTAACCAAAACCAAGTGCAG 60
DB 198 ATGATCGCGCTGCGATATGCGCTGCTTATTCACGSGAGCTAACCAAAACCAAGTGCAG 257
QY 61 AAGGTTGACCAATTTCTCTACACATGCGTCTCTCAGATGAGACCCCTTCTGGAGATTCT 120
DB 258 AAGGTTGACCAATTTCTCTACACATGCGTCTCTCAGATGAGACCCCTTCTGGAGATTCT 317
QY 121 AGCGCGTTCCGGAAGAGATGAGAAAGGGCTAGAGCTTACCAGCACCCTACAGAGCT 180
DB 318 AGCGCGTTCCGGAAGAGATGAGAAAGGGCTAGAGCTTACCAGCACCCTACAGAGCT 377
QY 181 GTGAATATGTTGGCTTACCTTTGTAGTCACTCCGAGTGGAGACCAATGGGAGTTTC 240
DB 378 GTGAATATGTTGGCTTACCTTTGTAGTCACTCCGAGTGGAGACCAATGGGAGTTTC 437
QY 241 CTGCGCTCGATCTTGAAGAACCAACTCCGTGCTCCGAGTAAAGGTGACGAGCAAT 300
DB 438 CTGCGCTCGATCTTGAAGAACCAACTCCGTGCTCCGAGTAAAGGTGACGAGCAAT 497
QY 301 GGCTCTCCAGAGATGAGATGAGAACCAAGATCTACGCCATCTTGAAGACATCATCGCG 360
DB 498 GGCTCTCCAGAGATGAGATGAGAACCAAGATCTACGCCATCTTGAAGACATCATCGCG 557
QY 361 GGCACTGGAACCCAGCTGTTTGAACCAATCGCGAATGCTGGCCAACTTCATGAGCAAG 420
DB 558 GGCACTGGAACCCAGCTGTTTGAACCAATCGCGAATGCTGGCCAACTTCATGAGCAAG 617
QY 618 CTACAAATCAAGAAGAGAGAGCTCCCTCGGGTTTACCTTCTCGTCCCTGCCACAG 677
DB 421 CTACAAATCAAGAAGAGAGAGCTCCCTCGGGTTTACCTTCTCGTCCCTGCCACAG 480
QY 481 ACAAACCTGAGATGAGATTTTGTGCTGCTGAGCTAAGGGGTTCAAGTCCAGTGGCGTG 540
DB 678 ACAAACCTGAGATGAGATTTTGTGCTGCTGAGCTAAGGGGTTCAAGTCCAGTGGCGTG 737
QY 541 GAAGGCAGAGATGTGTGAGCTGATCCGAGGTTTATCCAGCGAGAGGGAGCTTTGAC 600
DB 738 GAAGGCAGAGATGTGTGAGCTGATCCGAGGTTTATCCAGCGAGAGGGAGCTTTGAC 797
QY 601 ATTGACATTTGTGGCCGTGTGTAATGACACAGTTGGGACCATGATGACTTGGCTATGAT 660
DB 798 ATTGACATTTGTGGCCGTGTGTAATGACACAGTTGGGACCATGATGACTTGGCTATGAT 857
QY 661 GATCGAAGCTCGAGATTTGCTCATTTGTGGCACTGGCAGCAAGCCCTGCTACATGAG 720
DB 858 GATCGAAGCTCGAGATTTGCTCATTTGTGGCACTGGCAGCAAGCCCTGCTACATGAG 917
QY 721 GAATATGCGTCATATTGATGATGATGAGGAGATGAGGGGCGATGTCATCAACATGAG 780
DB 918 GAATATGCGTCATATTGATGATGATGAGGAGATGAGGGGCGATGTCATCAACATGAG 977
QY 781 TGGGAGAGCTTTGGGAGAGCGGTACACTCAATGACATCCGAACGAGTTTGAACGAGAG 840
DB 978 TGGGAGAGCTTTGGGAGAGCGGTACACTCAATGACATCCGAACGAGTTTGAACGAGAG 1037
QY 841 ATCGACATGGGCTCGCTGACACCTTGGAAGAGAGCTGTTTGAAGAGATGATTACGGGATG 900
DB 1038 ATCGACATGGGCTCGCTGACACCTTGGAAGAGAGCTGTTTGAAGAGATGATTACGGGATG 1097

QY 901 TTACATGGGAGAGCTGTGAGGCTCACTCTGTAAGATGGCCAGAGAGCTGTTGTTTC 960
DB 1098 TTACATGGGAGAGCTGTGAGGCTCACTCTGTAAGATGGCCAGAGAGCTGTTGTTTC 1157
QY 961 CAAGGAAACATCGCCAGCAACCTCTTACACTGCTCCCTTGGAGACCAAGATGTCTCG 1020
DB 1158 CAAGGAAACATCGCCAGCAACCTCTTACACTGCTCCCTTGGAGACCAAGATGTCTCG 1217
QY 1021 GATATTGAAGAGATTAAGATGGAATCGAAGAGGCTTACCAATCTCTGATGGCTGGGT 1080
DB 1218 GATATTGAAGAGATTAAGATGGAATCGAAGAGGCTTACCAATCTCTGATGGCTGGGT 1277
QY 1081 CTGAATTCATTGTCAGAGAGATTTGTGGCCAGCAGCAACCAATCTGCAGATTGTGTCACG 1140
DB 1278 CTGAATTCATTGTCAGAGAGATTTGTGGCCAGCAGCAACCAATCTGCAGATTGTGTCACG 1337
QY 1141 CGCTGGGCGAGTCTGTGGCCAGCACCCTGCGCGGGTCTGTGGCGAATCAAGAAGAAC 1200
DB 1338 CGCTGGGCGAGTCTGTGGCCAGCACCCTGCGCGGGTCTGTGGCGAATCAAGAAGAAC 1397
QY 1201 AAGGCGGAGAGAGAGCTTCCCTCCACCATCGGTGTGATGGCTCCGTCTCAAGAAGAAC 1260
DB 1398 AAGGCGGAGAGAGAGCTTCCCTCCACCATCGGTGTGATGGCTCCGTCTCAAGAAGAAC 1457
QY 1261 CCCCATTTTGGCCAGAGGCTCCCATTAAGGCAATGAGAGGCTGTGGCCGAGCTGTGATGTC 1320
DB 1458 CCCCATTTTGGCCAGAGGCTCCCATTAAGGCAATGAGAGGCTGTGGCCGAGCTGTGATGTC 1517
QY 1321 CGCTTCTCTCCGCTGTGAGATGGACGCGCAAGGGGGCTGCTATGATGACGGCGGTGGCT 1380
DB 1518 CGCTTCTCTCCGCTGTGAGATGGACGCGCAAGGGGGCTGCTATGATGACGGCGGTGGCT 1577
QY 1381 TTCCGCTTGGCTGACCAACACCGGGGCCCGCCAGAAAGACCTTGAAGTCTTGAAGCTGAGC 1440
DB 1578 TTCCGCTTGGCTGACCAACACCGGGGCCCGCCAGAAAGACCTTGAAGTCTTGAAGCTGAGC 1637
QY 1441 CAGAGAGAGCTTCTGAGAGTTAAGAGAAAGATGGAATGAGAGAGGAGCTGTGAGC 1500
DB 1638 CAGAGAGAGCTTCTGAGAGTTAAGAGAAAGATGGAATGAGAGAGGAGCTGTGAGC 1697
QY 1501 AAGGAGAGCATGCGGTGCGCCCTGTGAAGATGCTGCCACTTACGTGTGCTCCACTCA 1560
DB 1698 AAGGAGAGCATGCGGTGCGCCCTGTGAAGATGCTGCCACTTACGTGTGCTCCACTCA 1757
QY 1561 GATGGACAGAGAAAGAGAGACTTCTTGGCTTGGATCTTGGAGAAACAATCTCGGGGTC 1620
DB 1758 GATGGACAGAGAAAGAGAGACTTCTTGGCTTGGATCTTGGAGAAACAATCTCGGGGTC 1817
QY 1621 CTGCTGCTGCTGTGCTGATGGAAGCGAGGGGCTGGAATGATCAACAAGATCTTAC 1680
DB 1818 CTGCTGCTGCTGTGCTGATGGAAGCGAGGGGCTGGAATGATCAACAAGATCTTAC 1877
QY 1681 TCCATCCACAGAGAGGTTATGCAATGGACACTGGGGAGAGCTTCTGACCAATTTGTCAG 1740
DB 1878 TCCATCCACAGAGAGGTTATGCAATGGACACTGGGGAGAGAGCTTCTGACCAATTTGTCAG 1937
QY 1741 TGCATTGGCGAGCTTCTGTGAGTACATGGGCAATGAAGGGGCTGCTGCTTGGGTTTC 1800
DB 1938 TGCATTGGCGAGCTTCTGTGAGTACATGGGCAATGAAGGGGCTGCTGCTTGGGTTTC 1997
QY 1801 ACATTTCTCTTCCCTTGGCAGAGAAACAGCTAGACCAAGATCTCTCTCAAGTGGACA 1860
DB 1998 ACATTTCTCTTCCCTTGGCAGAGAAACAGCTAGACCAAGATCTCTCTCAAGTGGACA 2057
QY 1861 AAGGATTCAGAGCATTTGCTGCGAGGGTGAAGATGTGTCACCTTGTGAAGAAGCG 1920
DB 2058 AAGGATTCAGAGCATTTGCTGCGAGGGTGAAGATGTGTCACCTTGTGAAGAAGCG 2117
QY 1921 ATTCAACCGCGAGAGAGAGTTTACCTGGAATGTGTTGCCGTGTGTAATCAACAGTTGGG 1980
DB 2118 ATTCAACCGCGAGAGAGAGTTTACCTGGAATGTGTTGCCGTGTGTAATCAACAGTTGGG 2177

QY	1981	ACTATGATGACTTGTGGCTACGAGAACCTCTACTGTGAAGTTGGCTCATTTGTTGGCACC	2040
Db	2178	ACATATGATGACTTGTGGCTACGAGAACCTCTACTGTGAAGTTGGCTCATTTGTTGGCACC	2237
QY	2041	GGAAGCAGCCCTGCTACATGGAAGAGATGCTATATGTGAGCTGTGTGGACGGAGAGAG	2100
Db	2238	GGAAGCAGCCCTGCTACATGGAAGAGATGCTATATGTGAGCTGTGTGGACGGAGAGAG	2237
QY	2101	GGAAGCAGCTGTGTCAACATGAGATGGAGATTTGGGAGCAATTTGGCTGCTGATGAC	2160
Db	2298	GGAAGCAGCTGTGTCAACATGAGATGGAGATTTGGGAGCAATTTGGCTGCTGATGAC	2357
QY	2161	TTGGGAGCCCTGTTTGTATGTTGCTGTGTGATGAGCTTTCTCTCAACCTCGCAACAGAG	2220
Db	2358	TTGGGAGCCCTGTTTGTATGTTGCTGTGTGATGAGCTTTCTCTCAACCTCGCAACAGAG	2417
QY	2221	TTGGAGAGATGATCAGCGGCTACTTGGAGAGATTTGGGAGCAATTTGGCTGCTGATGAC	2280
Db	2418	TTGGAGAGATGATCAGCGGCTACTTGGAGAGATTTGGGAGCAATTTGGCTGCTGATGAC	2477
QY	2281	TTGACGAGCGGGGCTGCTCTTCCGAGCGGCTCTCAGAGCGCTCAAGACAAAGGGA	2340
Db	2478	TTGACGAGCGGGGCTGCTCTTCCGAGCGGCTCTCAGAGCGCTCAAGACAAAGGGA	2537
QY	2341	ATCTCTAAACTAAGTTCTCTCTCAATAGAGAGCAGCTGCTAGCCCTGTACAGGTT	2400
Db	2538	ATCTCTAAACTAAGTTCTCTCTCAATAGAGAGCAGCTGCTAGCCCTGTACAGGTT	2597
QY	2401	CGTGGCCTCTGGCGCCACTAGGCGTGGAGAGCAGTGGCAGTACAGCATCATGTGAG	2460
Db	2598	CGTGGCCTCTGGCGCCACTAGGCGTGGAGAGCAGTGGCAGTACAGCATCATGTGAG	2657
QY	2461	GAGGTGTGACTGTGTTGCCCGGCGCTGACAGCTCTGTGGCGAGGAGGAGCGGCC	2520
Db	2658	GAGGTGTGACTGTGTTGCCCGGCGCTGACAGCTCTGTGGCGAGGAGGAGGAGCGGCC	2717
QY	2521	GTAGTGACAAATGAAAGAGAACCGTGGGCTGGACAAACCCCAAGTGACAGTGGCGTG	2580
Db	2718	GTAGTGACAAATGAAAGAGAACCGTGGGCTGGACAAACCCCAAGTGACAGTGGCGTG	2777
QY	2581	GAGGAGCTCTGTATAGCTTCATCTCCTTGGCCAAAGTGCATGATGAGAGCGGTAGA	2640
Db	2778	GAGGAGCTCTGTATAGCTTCATCTCCTTGGCCAAAGTGCATGATGAGAGCGGTAGA	2837
QY	2641	GATCTGGCTCCGAATGTGACGTGTCTTCTGGAATCCAGAGACGGCAGTGGGAGGGA	2700
Db	2838	GATCTGGCTCCGAATGTGACGTGTCTTCTGGAATCCAGAGACGGCAGTGGGAGGGA	2897
QY	2701	GCAAGCTCTCATCTGCGCGGCTGCGGCTGCGGAGGCTGGGAGAGATAG	2754
Db	2898	GCAAGCTCTCATCTGCGCGGCTGCGGCTGCGGAGGCTGGGAGAGATAG	2951
RESULT 4			
LOCUS	RATHKII	3635 bp	mRNA linear ROD 27-APR-1993
DEFINITION	Rat hexokinase type II (HKII) mRNA, complete cds.		
ACCESSION	M68971.1	GI:204612	
VERSION	M68971.1	GI:204612	
KEYWORDS	hexokinase type II.		
SOURCE	Rat (strain C.R. sd)	skeletal (soleus) muscle, cDNA to mRNA.	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 3635)		
AUTHORS	Thelen, A.P. and Willson, J.E.		
TITLE	Complete amino acid sequence of the type II isozyme of rat hexokinase, deduced from the cloned cDNA: comparison with a hexokinase from novikoff ascites tumor		
JOURNAL	Arch. Biochem. Biophys. 286 (2), 645-651 (1991)		
MEDLINE	91378366		
PUBMED	1897984		

FEATURES	Location/Qualifiers		
source	1..3635		
	/organism="Rattus norvegicus"		
	/strain="C.R. sd"		
	/db_xref="taxon:10116"		
	/tissue_type="Soleus muscle"		
	/tissue_lib="lambda-gt10"		
gene	1..3635		
	/gene="HKII"		
	/gene="2951"		
CDS	/EC_number="2.7.1.1"		
	/codon_start=1		
	/product="hexokinase type II"		
	/protein_id="AA41333.1"		
	/db_xref="GI:204613"		
	/translation="MTASHIACLFTELNOYOKVDLFYHRLSDFTLEISRRF KEMKIGATHTPTAAVKMLPTFVRSPDTEHGFALDLGGTNRFLVRVTDNGL QVEMENOIYAIPEIDIMRGSGTQDFHIAECLANFMDKIDIKERKPLGTFSPCHQ TKLDESFLVMTKGFSSGVEGDDVDLIRKAIORRDEPIDIAVAVNDVGTMTG YDQONCEIGLIVGTGSNACVMEERHIDVVEGEGRMCIIMWGAFGDDGTNDIRE FDREIDMSLPGKOLFEMKISGMVGEIYRLILYKAKKELLFOGLSPELLTTGSF ETRDVSDIEDKXGIEKAXOILIRLGLNPQEDCVATHRCQIYSTRSASLCATTLA VLRIRKNEKEERURSTIGVDSYKKNPFAKRLKRAVRLVPDQVRLKESDGS KGAAWTAVALYRLADQHRAROKTLESILKHELEVRKMKVMEGSLKETHAVAP VKMLPTVYCATPDTGKDFLADLGGTNRFLVLRNKRGRVEMHNTYSIPQV MHTGEELPDHIVQCLADFLVGMKVSPLFTFSPQOQSLDSILKMTKGR ASGCEGDDVYTLKEATHRREEDLDVAVAVNDVGTMTGVEDEPHCEGLVGTGS NACVMEERHVEIDVEGEGRMCIIMWGAFGDDGTNDIREFDREIDMSLPGKOLF FEMKISGMVGEIYRLILIDFTKRLRGRISERLRTRIETFKLSQIESCLALL QVAILRHLDLESTCDSTLIVKEVCTVYARRAOLDCAGAAVADKIRENGIDNLK TVGVDTLTKLHPHFAKVMHEVYRDLAPKCDVSFLSEDSGGAALITVAACIRREA GOR"		
BASE COUNT	816 a 875 c 1098 g 846 t		
ORIGIN			
Query Match	92.6%; Score 2550; DB 10; Length 3635;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 2750; Conservative	0; Mismatches 4; Indels 0; Gaps 0;		
QY	1	ATGATGCGCTGGCATATGATGCGCTGTTATTCAGGAGCTCAACCAAAACCAAGTGAG	60
Db	198	ATGATGCGCTGGCATATGATGCGCTGTTATTCAGGAGCTCAACCAAAACCAAGTGAG	257
QY	61	AAGGTTACCAATTTCTCTACACATGCGTCTCAGATGAGACCTCTTGAGATTTCT	120
Db	258	AAGGTTACCAATTTCTCTACACATGCGTCTCAGATGAGACCTCTTGAGATTTCT	317
QY	121	AGCGGTTCCGGAAGAGATGGAAGAGGCTAGAGCTACACGACCTACAGCAGCT	180
Db	318	AGCGGTTCCGGAAGAGATGGAAGAGGCTAGAGCTACACGACCTACAGCAGCT	377
QY	181	GTGAAATGTTGCTTACCTTGTGAGTGCAACTCCGATGGAGGAGCAAGATGGAGGTT	240
Db	378	GTGAAATGTTGCTTACCTTGTGAGTGCAACTCCGATGGAGGAGCAAGATGGAGGTT	437
QY	241	CTGGCTCTGATCTTTGGAGAACCACTCCGTGCTCGAGTAAGAGGTGACGAGCAAT	300
Db	438	CTGGCTCTGATCTTTGGAGAACCACTCCGTGCTCGAGTAAGAGGTGACGAGCAAT	497
QY	301	GGCCTCAGAGAGTGGAGATGGAACCAAGTCTACGCCATCTTGAGGAGCATGATCGG	360
Db	498	GGCCTCAGAGAGTGGAGATGGAACCAAGTCTACGCCATCTTGAGGAGCATGATCGG	557
QY	361	GGCAGTGAACCCAGCTGTTTGAACCATGCGCGAATGCGCTGGCACTTATGAGCAAG	420
Db	558	GGCAGTGAACCCAGCTGTTTGAACCATGCGCGAATGCGCTGGCACTTATGAGCAAG	617
QY	421	CTACAAATCAAGAGAAAGAGCTCCCTCTGGGTTTCACTTCTGTTCCCTGGCACAG	480
Db	618	CTACAAATCAAGAGAAAGAGCTCCCTCTGGGTTTCACTTCTGTTCCCTGGCACAG	677
QY	481	ACAAACTGATGAGAGTTTGTGCTCGTGACCTAAGGAGGTTCAAGTCCAGTGCGGTG	540

|||||
Db 678 AAAAACTGATGAGAGTTTTTGTCTGTGGACTAAGGGGTTCAAGTCCAGTGCCTG 737
OY 541 GAAGGACAGATGTGTGACCTGATCCGGAAGGTTATCCAGCGCAGAGGACTTTGAC 600
Db 738 GAAGGACAGATGTGTGACCTGATCCGGAAGGTTATCCAGCGCAGAGGACTTTGAC 797
OY 601 ATTGACATTTGGCCGTGTGTAATGACAGATTTGGACCATGATGATGATGATGAT 660
Db 798 ATTGACATTTGGCCGTGTGTAATGACAGATTTGGACCATGATGATGATGATGAT 857
OY 661 GATCAGAACTGCGAGATTGCTCATTTGGCACTGCGACAAAGCCCTGCTACATGAG 720
Db 858 GATCAGAACTGCGAGATTGCTCATTTGGCACTGCGACAAAGCCCTGCTACATGAG 917
OY 721 GAAATGCGCATATTGACATTTGGGAGGAGATGAGGGGCGATGTCATCAACATGAG 780
Db 918 GAAATGCGCATATTGACATTTGGGAGGAGATGAGGGGCGATGTCATCAACATGAG 977
OY 781 TGGGAGACCTTTGGGAGCAGCAGTACACTCAATGACATCCGAACGAGTTTGACCGAG 840
Db 978 TGGGAGACCTTTGGGAGCAGCAGTACACTCAATGACATCCGAACGAGTTTGACCGAG 1037
OY 841 ATCGACATGGCTCGCTGACACCTTGGGAAGCAGCTGTTTGAAGAATGATTAAGCGGATG 900
Db 1038 ATCGACATGGCTCGCTGACACCTTGGGAAGCAGCTGTTTGAAGAATGATTAAGCGGATG 1097
OY 901 TACATGAGGGGAGCTGGTGCAGCTCCGTCATCCGTAAGATGGCCAAAGCAGATGTTGTC 960
Db 1098 TACATGAGGGGAGCTGGTGCAGCTCCGTCATCCGTAAGATGGCCAAAGCAGATGTTGTC 1157
OY 961 CAAGGAAACTCAGCCCAAGCAACCTCTTACACTGCTCCTTCGAGCAGCAAAAGATGCTCG 1020
Db 1158 CAAGGAAACTCAGCCCAAGCAACCTCTTACACTGCTCCTTCGAGCAGCAAAAGATGCTCG 1217
OY 1021 GATATTGAAGAGATTAAGATGAATCGAAGAGGCTACCAATCTGATGCGCTGGGT 1080
Db 1218 GATATTGAAGAGATTAAGATGAATCGAAGAGGCTACCAATCTGATGCGCTGGGT 1277
OY 1081 CTGAATCCATTTCAGAGAGATTGTGTGGCAGCAGCAACGATTCGACAGATTTGTCACAG 1140
Db 1278 CTGAATCCATTTCAGAGAGATTGTGTGGCAGCAGCAACGATTCGACAGATTTGTCACAG 1337
OY 1141 CGCTGGCCAGTCTGTGGCAGCAGCAACCTTGGCCGCGTCTGTGGCAATCAAGAAGAAC 1200
Db 1338 CGCTGGCCAGTCTGTGGCAGCAGCAACCTTGGCCGCGTCTGTGGCAATCAAGAAGAAC 1397
OY 1201 AAGGCGAGAGAGCACTTCCGTCACACCATGGTGTGATGGCTCCGTCACAAAGAACAT 1260
Db 1398 AAGGCGAGAGAGCACTTCCGTCACACCATGGTGTGATGGCTCCGTCACAAAGAACAT 1457
OY 1261 CCCCATTTCAGCAAGCGTTCATTAAGGAGTGAAGAGGCTGTGGCCGACAGTGTATGTC 1320
Db 1458 CCCCATTTCAGCAAGCGTTCATTAAGGAGTGAAGAGGCTGTGGCCGACAGTGTATGTC 1517
OY 1321 CGCTTCCTCCGCTGTGAGAGTGGACAGCGGCAAGGGGCTGTATGGTGAAGCGGTGCT 1380
Db 1518 CGCTTCCTCCGCTGTGAGAGTGGACAGCGGCAAGGGGCTGTATGGTGAAGCGGTGCT 1577
OY 1381 TACCGTGTGGCTGACCAACACCGGGGCCCCGCAAGAACCTTGGATGTTGAAGCTGAGAC 1440
Db 1578 TACCGTGTGGCTGACCAACACCGGGGCCCCGCAAGAACCTTGGATGTTGAAGCTGAGAC 1637
OY 1441 CACGAGCAGCTTGTGAGGTTTAAGAGAAGATGAAGGTGAAGAGAGGAGGCTGAGAC 1500
Db 1638 CACGAGCAGCTTGTGAGGTTTAAGAGAAGATGAAGGTGAAGAGAGGAGGCTGAGAC 1697
OY 1501 AAGGAGCAGCATGGGTGGCCCCCTGTGAAGATGTCGCCACTTACGTGTGTCCACTCCA 1560
Db 1698 AAGGAGCAGCATGGGTGGCCCCCTGTGAAGATGTCGCCACTTACGTGTGTGTCCACTCCA 1757
OY 1561 GATGGCAGAGAAAGAGACCTTGTGGCTTGGATCTTGGAGAACAACTTCCGGGTC 1620
|||||

Db 1758 GATGGCAGAGAAAGAGACCTTCTTGCCCTTGGATCTTGAAGAACAACTTCCGGGTC 1817
OY 1621 CTGCTGTGCTGTGTGTGCTAATGGCAAGCGGAGGGGCTGTGAAGTGTATTAACAAGATCTAC 1680
Db 1818 CTGCTGTGCTGTGTGTGCTAATGGCAAGCGGAGGGGCTGTGAAGTGTATTAACAAGATCTAC 1877
OY 1681 TCCATCCCAAGAGAGATTATGATGGCACTGGGGAAGAGCTCTTGACACAAATTGGCCAG 1740
Db 1878 TCCATCCCAAGAGAGATTATGATGGCACTGGGGAAGAGACTCTTGACACCAATTTGCCAG 1937
OY 1741 TGCATTCGCGACTTCTGTGAGTACATGGCATGAAGGGGCTGTCCCTGCTTTGGGTTTC 1800
Db 1938 TGCATTCGCGACTTCTGTGAGTACATGGCATGAAGGGGCTGTCCCTGCTTTGGGTTTC 1997
OY 1801 ACATTCCTCTCCCTTGGCAGAGAAACAGCTTACACAGATCTCTCTCAAGTGGANA 1860
Db 1998 ACATTCCTCTCCCTTGGCAGAGAAACAGCTTACACAGATCTCTCTCAAGTGGANA 2057
OY 1861 AAGGATTCAGGCAATCTGGCTGCGAGGGTGAAGATGTGTACCTTCTGTAAGAAAGCA 1920
Db 2058 AAGGATTCAGGCAATCTGGCTGCGAGGGTGAAGATGTGTACCTTCTGTAAGAAAGCA 2117
OY 1921 ATTCAACGCGAGAGAGATTGATCACTGATGTGTTGCCGTGTGATGACACAGTTGGG 1980
Db 2118 ATTCAACGCGAGAGAGATTGATCACTGATGTGTTGCCGTGTGATGACACAGTTGGG 2177
OY 1981 ACTATGATGACTTGTGGCTACGAAGACCTCTACTGTAAGTGGCTCATTTGTGGCAC 2040
Db 2178 ACTATGATGACTTGTGGCTACGAAGACCTCTACTGTAAGTGGCTCATTTGTGGCAC 2237
OY 2041 GGAAGCAACGCTTGTACATGGAAGAGATGCTTAATGTGAGACTGTGTGAGAGAGAGAG 2100
Db 2238 GGAAGCAACGCTTGTACATGGAAGAGATGCTTAATGTGAGACTGTGTGAGAGAGAGAG 2297
OY 2101 GGACGATGTGTGTCAACATGAGATGGGAGCATTTGGGGAACAATGTGCTGTGATGAC 2160
Db 2298 GGACGATGTGTGTCAACATGAGATGGGAGCATTTGGGGAACAATGTGCTGTGATGAC 2357
OY 2161 TTGCGGACCGCTTGTGATTTGCTGTGATGAGCTTCTCAACCTGGCAAAACAGAG 2220
Db 2358 TTGCGGACCGCTTGTGATTTGCTGTGATGAGCTTCTCAACCTGGCAAAACAGAG 2417
OY 2221 TTGCAAGATGATCAGCGGATGATCTTGGAGAGATTTGTGCGCAACTTCTCATCAT 2280
Db 2418 TTGCAAGATGATCAGCGGATGATCTTGGAGAGATTTGTGCGCAACTTCTCATCAT 2477
OY 2281 TTCACGAAGCGGGGCTGTCTTCCGAGGGCCCATCTGACAGCGCTCAAGAACAGGGA 2340
Db 2478 TTCACGAAGCGGGGCTGTCTTCCGAGGGCCCATCTGACAGCGCTCAAGAACAGGGA 2537
OY 2341 ATCTCTGAACATTAAGTTCCTGCTCAGATGAGAGAGCACTGCTAGCCCTGCTACAGTT 2400
Db 2538 ATCTCTGAACATTAAGTTCCTGCTCAGATGAGAGAGCACTGCTAGCCCTGCTACAGTT 2597
OY 2401 CGTGCCATCTCGCCACCTAGGGCTGAGAGCAGCTGCGATGACAGCATCATCTGTGAAG 2460
Db 2598 CGTGCCATCTCGCCACCTAGGGCTGAGAGCAGCTGCGATGACAGCATCATCTGTGAAG 2657
OY 2461 GAGCTGTGACAGTGTGTGTTCCCGGCGCGCTGACAGACCTTGTGGGCGAGCATGTGCCG 2520
Db 2658 GAGCTGTGACAGTGTGTGTTCCCGGCGCGCTGACAGACCTTGTGGGCGAGCATGTGCCG 2717
OY 2521 GTAGTGAACAAATTAAGAGAAACCGTGGCTGGAACAACCCCAAGTGAACATGTGGCGTG 2580
Db 2718 GTAGTGAACAAATTAAGAGAAACCGTGGCTGGAACAACCCCAAGTGAACATGTGGCGTG 2777
OY 2581 GACGGGACTTGTATTAAGCTTCACTTGTGCCAAGTCAATGAGAGCGGTGAGA 2640
Db 2778 GACGGGACTTGTATTAAGCTTCACTTGTGCCAAGTCAATGAGAGCGGTGAGA 2837
OY 2641 GATCTGGCTCCGAATGTGAGCTGTCTTCCGGAATCCGAGAGCGGAGTGGGAAGGA 2700
Db 2838 GATCTGGCTCCGAATGTGAGCTGTCTTCCGGAATCCGAGAGCGGAGTGGGAAGGA 2897
|||||

RESULT 5
RATHEKIIINH
LOCUS Rat hexokinase II (HKII) mRNA, 3' end.
DEFINITION M68972
ACCESSION M68972
VERSION M68972.1 GI:204614
KEYWORDS hexokinase type II.
SOURCE Rat Novikoff ascites hepatoma, CDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
Rattus.
Rattus.
1 (bases 1 to 2236)
REFERENCE Thelen,A.P. and Wilson,J.E.
AUTHORS Complete amino acid sequence of the type II isozyme of rat
TITLE hexokinase, deduced from the cloned cDNA: comparison with a
hexokinase from novikoff ascites tumor
JOURNAL Arch. Biochem. Biophys. 286 (2), 645-651 (1991)
MEDLINE 91378386
PUBMED 1897984
FEATURES
source location/Qualifiers
1..2236 /organism="Rattus norvegicus"
/db_xref="taxon:10116"
/lssue_type="novikoff ascites hepatoma"
1..2236 /gene="HKII"
<1..1552 /gene="HKII"
/codon_start=2 /product="hexokinase type II"
/protein_id="AA01354.1" /db_xref="GI:204615"
/translation="GEEERLSTIGVDGVYKKHHPFAKRLHKAARLVLPDQVFLRS
EDSGGAAMAVTAVAYRLADHAROKTLESLKSHOLEVRKRMKEVMEQSLKET
HAYAPVKMLPTVYCATPPDREKGDPLADIGTNPRLIYRVNGKRGRVEMHNKITS
IIPVNHGTSEELFDHIVOCIAIDPLEYLMGKGVSLPVTGTFSPCCQNSIDQSLILK
TKGFKASGCEGEVDVTLKEAIRREFEEDVVAVVNDVGTMTGCEYEDPHEVGLI
VGTGSNACVMEKRWELVDGEGRMCVNMENGAFGNGDLDRITFVDAVELSLN
PGKOREKMSIEMYLGEIYRNILIDFTKRLFRGRISERLKTGIFETVFLSQIESD
CLALLOYRALIRHIGLESTCDSDIIVKEVTVVARRAOCGAGMAVVVDKIRENRL
DNLKVTGVVDGTLVXKLHPHFAKVMHETVRDLARKQDVSFLESDGSKGALITAVAC
RIRDAQR"
BASE COUNT 485 a 524 c 680 g 547 t
ORIGIN
Query Match 52.7%; Score 1450; DB 10; Length 2236;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1550; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1203 GGGGAGAGGAGGACTTGGCTCCACCATCGGTGATGGCTCGCTACAAGAAAACATCC 1262
DB 1 GGGGAGAGGAGGACTTGGCTCCACCATCGGTGATGGCTCGCTACAAGAAAACATCC 60
QY 1263 CCATTTTCCCAAGGCTCTCCATAAGGAGTGAAGAGGTGTGCCGACTGTGATGTCCG 1322
DB 61 CCATTTTCCCAAGGCTCTCCATAAGGAGTGAAGAGGTGTGCCGACTGTGATGTCCG 120
QY 1323 CTTCCTCGGCTGTAGATGAGTGAAGGAGGAGGAGGCTATGAGTGAAGGAGGAGCTTA 1382
DB 121 CTTCCTCGGCTGTAGATGAGTGAAGGAGGAGGAGGAGGCTATGAGTGAAGGAGCTTA 180
QY 1383 CCGTCTCGGCTGAACACACCGGGCCCGCAGAGACCTTGAGTCTCTGAAGCTGAGCCA 1442
DB 181 CCGTCTCGGCTGAACACACCGGGCCCGCAGAGACCTTGAGTCTCTGAAGCTGAGCCA 240
QY 1443 CGAGCAGCTTCTGAGAGTTAAGACAGATGAAGTGAATGAGCAGGCTCTGAGCAA 1502

DB 241 CGAGCAGCTTCTGAGAGTTAAGACAGATGAAGTGAATGAGAAATGAGAGGCTCTGAGCAA 300
QY 1503 GAGAGCGATCGGGTGGCCCGCTGTGAAGATGTGCGCCACTACGNTGTGCACCTCCAGA 1562
DB 301 GAGAGCGATCGGGTGGCCCGCTGTGAAGATGTGCGCCACTACGNTGTGCACCTCCAGA 360
QY 1563 TGGCAGACAGAAAGAGACTTCTTGACCTTGATCTTGAGAAACAACCTTCGGGTCCT 1622
DB 361 TGGCAGACAGAAAGAGACTTCTTGACCTTGATCTTGAGAAACAACCTTCGGGTCCT 420
QY 1623 GCTGTGCTGTGCGTAATGAGCAAGCGAGGCGGTGAGATGCATTAACAGATCTACTC 1682
DB 421 GCTGTGCTGTGCGTAATGAGCAAGCGAGGCGGTGAGATGCATTAACAGATCTACTC 480
QY 1683 CATCCACAGAGAGTTATGATGATGACACTGGGGAAGAGCTCTTCGACCAATTCGCAATG 1742
DB 481 CATCCACAGAGAGTTATGATGATGACACTGGGGAAGAGCTCTTCGACCAATTCGCAATG 540
QY 1743 CATTGCGGACTTCTGAGTACATGAGCATGAAGGCGCTGCCCTTGGGTTTCAC 1802
DB 541 CATTGCGGACTTCTGAGTACATGAGCATGAAGGCGCTGCCCTTGGGTTTCAC 600
QY 1803 ATTCTCTTCCCTTGGCCAGCAGAACAGCTTAGACCAGCATCTCTCAAGTGACAAA 1862
DB 601 ATTCTCTTCCCTTGGCCAGCAGAACAGCTTAGACCAGCATCTCTCAAGTGACAAA 660
QY 1863 GGGATTCAGGCACTTGCTGCTCGAAGGAGGATGTGTCACCTTGCTGAAGAAAGCAT 1922
DB 661 GGGATTCAGGCACTTGCTGCTCGAAGGAGGATGTGTCACCTTGCTGAAGAAAGCAT 720
QY 1923 TCACCGGCGAGAGAGATTGACCTGGATGTGTCCTGGTGAATGACACAGTTGGGAC 1982
DB 721 TCACCGGCGAGAGAGATTGACCTGGATGTGTCCTGGTGAATGACACAGTTGGGAC 780
QY 1983 TATGATGACTTGTGGCTACGAAGACCTCAGTGTGAATGTGGCTCATTTGTGGACCGG 2042
DB 781 TATGATGACTTGTGGCTACGAAGACCTCAGTGTGAATGTGGCTCATTTGTGGACCGG 840
QY 2043 AAGCAACCCCTGCTACATGAGAAAGATCGTAAATGTGAGCTGTGAGCAGGAGAGAGG 2102
DB 841 AAGCAACCCCTGCTACATGAGAAAGATCGTAAATGTGAGCTGTGAGCAGGAGAGAGG 900
QY 2103 ACGATGTGTCAACATGAGAGTGGGAGCATTTGGGAGCAATGGCTCGTGATGACTT 2162
DB 901 ACGATGTGTCAACATGAGAGTGGGAGCATTTGGGAGCAATGGCTCGTGATGACTT 960
QY 2163 GCGGACGCTGTGATGTGTGTGATGATGAGTGTCTCTCAACCTGGCAAAACAGAGTT 2222
DB 961 GCGGACGCTGTGATGTGTGTGATGATGAGTGTCTCTCAACCTGGCAAAACAGAGTT 1020
QY 2223 CGAAGATGATCAGCGGATGATCTTGGGAGAGATTGTGCGCAACATCTCATGATTT 2282
DB 1021 CGAAGATGATCAGCGGATGATCTTGGGAGAGATTGTGCGCAACATCTCATGATTT 1080
QY 2283 CACGAGCGGGGCGTCTCTTCGAGGCCGATCTCAGAGCGCTCAAGACAAAGGGGAT 2342
DB 1081 CACGAGCGGGGCGTCTCTTCGAGGCCGATCTCAGAGCGCTCAAGACAAAGGGGAT 1140
QY 2343 CTCTGAACCTAAGTTCCTGCTCAGATGAGAGGAGCTGTACCGCTGACAGAGTTG 2402
DB 1141 CTCTGAACCTAAGTTCCTGCTCAGATGAGAGGAGCTGTACCGCTGACAGAGTTG 1200
QY 2403 TGGCATCTGCGCCACCTAGAGGCTGAGAGACAGTGCATGACAGCATCTGTAAGA 2462
DB 1201 TGGCATCTGCGCCACCTAGAGGCTGAGAGACAGTGCATGACAGCATCTGTAAGA 1260
QY 2463 GGTGTGCACTGTGTTGCGCGCGCGCTGCAAGCTGTGTGGCGAAGCATGGCCCGCT 2522
DB 1261 GGTGTGCACTGTGTTGCGCGCGCGCTGCAAGCTGTGTGGCGAAGCATGGCCCGCT 1320
QY 2523 AGTGACAAAGATTAAGAGAAACGTTGGGCTGACAAACCCAAAGTGAATGTGGGCGTGA 2582

Db 1321 AGTGACAGAGATAGAGAGACCGTGGGCTGGACAACTCAAGTACAGTGGCGCTGGA 1380
 Oy 2583 CGGAGCTCTGTATACCTATCCTACCTTGTGGCAAGTCATGATGAGACGGTAGAGA 2642
 Db 1381 CGGAGCTCTGTATACCTATCCTACCTTGTGGCAAGTCATGATGAGACGGTAGAGA 1440
 Oy 2643 TCTGGCTCGGAATGTGACGTGCTTCTGATCCGAGCGGCGGCTATGGAGGAGC 2702
 Db 1441 TCTGGCTCGGAATGTGACGTGCTTCTGATCCGAGCGGCGGCTATGGAGGAGC 1500
 Oy 2703 ACTCTCATCATCTCGGCTGGCTGCGGCATCCGAGGCGCTGGAGAGATAG 2754
 Db 1501 ACTCTCATCATCTCGGCTGGCTGCGGCATCCGAGGCGCTGGAGAGATAG 1552

RESULT 6
 AC098769 231777 bp DNA linear HTG 13-JUL-2002
 LOCUS Rattus norvegicus clone CH230-186P10, *** SEQUENCING IN PROGRESS
 DEFINITION *** 16 unordered pieces.
 AC098769
 VERSION AC098769.10 GI:21735640
 KEYWORDS HTG: HTGS_PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
 AUTHORS 1 (bases 1 to 231777)
 Muzny,D.M., Adams,C., Adio-Oduola,B., All-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T., Barbieri,J., Benton,J., Blinge,K., Blankenburg,K., Bonaldi,D., Bouck,J., Bowls,S., Brleya,M., Brown,B., Brown,M., Bryant,N.P., Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Kralovic,U., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Lloisged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Maswhey,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Monabadi,K., Morgan,M., Morris,S., Nguyen,N., Nickerson,E., Nwokwenwo,S., Ogih,M., Okumura,G., Oragunye,N., Ovielo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Severy,G., Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I., Soderren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S., Williams,G., Williamson,A., Wleczek,R., Woden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Direct Submission
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 231777)
 REFERENCE
 AUTHORS Worley,K.C.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (01-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 231777)
 Worley,K.C.
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:20467513.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: G1SF
 Center clone name: CH230-186P10
 ----- Summary Statistics
 Sequencing vector: pBlasmid
 Chemistry: Dye-terminator Big Dye 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 216012 bases at least Q40
 Consensus quality: 217712 bases at least Q30
 Consensus quality: 218875 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preceeded.
 *
 * 1
 * 1531: contig of 1530 bp in length
 * 1531: gap of unknown length
 * 1631: contig of 1107 bp in length
 * 2738: gap of unknown length
 * 2838: contig of 1515 bp in length
 * 4353: gap of unknown length
 * 4453: contig of 1106 bp in length
 * 5559: gap of unknown length
 * 5659: contig of 1283 bp in length
 * 6842: gap of unknown length
 * 7042: contig of 2672 bp in length
 * 9714: gap of unknown length
 * 9813: contig of 10465 bp in length
 * 9814: gap of unknown length
 * 20279: contig of 8044 bp in length
 * 20379: gap of unknown length
 * 28423: contig of 9275 bp in length
 * 28523: gap of unknown length
 * 37798: contig of 16661 bp in length
 * 37898: gap of unknown length
 * 54559: contig of 18169 bp in length
 * 54659: gap of unknown length
 * 72828: contig of 27423 bp in length
 * 72928: gap of unknown length
 * 100351: contig of 26677 bp in length
 * 100450: gap of unknown length
 * 127128: contig of 28460 bp in length
 * 127228: gap of unknown length
 * 155687: contig of 33350 bp in length
 * 155788: gap of unknown length
 * 189137: contig of 42540 bp in length
 * 189238: gap of unknown length
 * 231777: contig of 42540 bp in length.
 Location/Qualifiers
 1..231777
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /clone="CH230-186P10"

BASE COUNT 60544 a 52226 c 52675 g 62183 t 4149 others
 ORIGIN
 Query Match 11.1% Score 305; DB 2; Length 231777;
 Best Local Similarity 100.0%; Pred. No. 9,6e-152;
 Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC098769 231777 bp DNA linear HTG 13-JUL-2002
 LOCUS Rattus norvegicus clone CH230-186P10, *** SEQUENCING IN PROGRESS
 DEFINITION ***16 unordered pieces.
 AC098769
 VERSION AC098769.10 GI:21735640
 KEYWORDS HTG; HTGS_PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 231777)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbarelle,J., Benton,J., Blinze,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthett,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Faltus,T., Ferraguto,D., Flagg,N., Ford,J., Foster,F., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gottlieb,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
 Homsl,F., Howard,S., Huber,J., Hui,Y., Hui,Y., Hummel,S.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kraviec,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Licharge,O., Liu,C., Liu,J., Liu,M., Liu,M.,
 Lozato,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okunolu,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
 Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
 Sodergren,E., Sonike,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalobos,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G., and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 231777)
 Worley,K.C.
 Direct Submission
 Submitted (01-NOV-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 231777)
 Worley,K.C.
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:20467513.

----- Genome Center
 Center code: BCM
 Center: Baylor College of Medicine
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: G1SF
 Center clone name: CH230-186P10
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 216012 bases at least Q40
 Consensus quality: 217712 bases at least Q30
 Consensus quality: 218875 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1	1530: contig of 1530 bp in length
1531	1630: gap of unknown length
1631	2737: contig of 1107 bp in length
2738	2837: gap of unknown length
2838	4352: contig of 1515 bp in length
4353	4452: gap of unknown length
4453	5558: contig of 1106 bp in length
5559	5658: gap of unknown length
5659	6941: contig of 1283 bp in length
6942	7041: gap of unknown length
7042	9713: contig of 2672 bp in length
9714	9813: gap of unknown length
9814	20278: contig of 10465 bp in length
20279	20378: gap of unknown length
20379	28422: contig of 8044 bp in length
28423	28522: gap of unknown length
28523	37797: contig of 9275 bp in length
37798	37897: gap of unknown length
37898	54558: contig of 16661 bp in length

polyA_signal	ACG# 11668 at locations 27179..27212"
5431..5436	
/gene="HKII"	
/note="Primary signal"	
/evidence=experimental	
5453	
/gene="HKII"	
/evidence=experimental	
5454..5474	
/note="Partial polyA tail"	
BASE COUNT	1320 a 1268 c 1561 g 1325 t
ORIGIN	
Query Match	3.4% Score 95; DB 10; Length 5474;
Best Local Similarity	99.0%; Pred. No. 6.5e-39;
Matches 195; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
OY 1246	GCTTACAGAAACATCCCATTTTGTCCAAAGCGCTTCATTAAGCAGTAGAGAGGCTGTG 1305
Db 1663	GTCCTACAGAAACATCCCATTTTGTCCAAAGCGCTTCATTAAGCAGTAGAGAGGCTGTG 1722
OY 1306	CCCGACTGTGATGTGCCGCTTCTCCGCGCTGTGAGATGACAGCGGCAAGGGGCTGTATG 1365
Db 1723	CCCGATTTGTGATGTCCGCTTCTCCGCGCTGTGAGATGACAGCGGCAAGGGGCTGTATG 1782
OY 1366	GTGACGGCGGTGGCTTACCGTCTGGCTGACCAACACCGGGCCCGCCAGAGACCTGTGAG 1425
Db 1783	GTGACGGCGGTGGCTTACCGTCTGGCTGACCAACACCGGGCCCGCCAGAGACCTGTGAG 1842
OY 1426	TCTCTGAAGCTGAGCCA 1442
Db 1843	TCTCTGAAGCTGAGCCA 1859
RESULT 9	
LOCUS	MWHEX113 28646 bp DNA linear ROD 02-MAR-2000
DEFINITION	Mus musculus gene for hexokinase II, exon 3-18.
ACCESSION	Y11668
VERSION	Y11668.1 GI:1907085
KEYWORDS	hexokinase II.
SOURCE	Mus musculus.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	Heikkinen, S., Supola, S., Malkki, M., Deeb, S.S., Janne, J. and Laakso, M.
TITLE	Mouse hexokinase II gene: structure, cDNA, promoter analysis, and expression pattern
JOURNAL	Mamm. Genome 11 (2), 91-96 (2000)
MEDLINE	20122158
PUBMED	10656921
REFERENCE	2 (bases 1 to 28646)
AUTHORS	Heikkinen, S.
TITLE	Direct Submission
JOURNAL	Submitted (06-MAR-1997) S. Heikkinen, University of Kuopio, A.I.Virtanen Institute, PO Box 1627, SF-70211 Kuopio, FINLAND
FEATURES	Location/Qualifiers
source	1..28646
	/organism="Mus musculus"
	/strain="129/SvJ"
	/db_xref="taxon:10090"
	/sex="female"
	/tissue_type="liver"
	/clone_id="lambda fixII"
	<1..2983
	/number=2
	1992..2187
	/partial
	/rpt_family="SINE/B2"
	/rpt_type=INVERTED
	2699..2830
	repeat_region

exon	/rpt_type=DIRECT /rpt_unit=2699. .2702 2984. .3132 /product="hexokinase II" /number=3 /used_in=y111666:hexIIrna 3133. .7348
intron	/number=3 /used_in=y111666:hexIIrna 3133. .7348
repeat_region	/number=3 4446. .4497 /rpt_type=DIRECT /rpt_unit=4446. .4447 6402. .6441
repeat_region	/rpt_type=DIRECT /rpt_unit=6402. .6403 7349. .7468 /product="hexokinase II" /number=4 /used_in=y111666:hexIIrna 7469. .8210 7469. .8210 /number=4 7920. .7959
intron	/rpt_type=DIRECT /rpt_unit=7920. .7921 8211. .8306
exon	/product="hexokinase II" /number=5 /used_in=y111666:hexIIrna /used_in=y111666:hexIICDS 8307. .8882 /number=5 8883. .8982
intron	/product="hexokinase II" /number=6 /used_in=y111666:hexIIrna /used_in=y111666:hexIICDS 8983. .9233 /number=6 9234. .9417
exon	/product="hexokinase II" /number=7 /used_in=y111666:hexIIrna /used_in=y111666:hexIICDS 9418. .12798 /number=7 9668. .9789
intron	/partial /note="B1-F-repeat" /rpt_family="SINE/Alu" /rpt_type=INVERTED 11156. .11299
repeat_region	/partial /note="B1-MM-repeat" /rpt_family="SINE/Alu" /rpt_type=INVERTED 11894. .12094
repeat_region	/partial /note="L1-MM-repeat" /rpt_family="LINE/L1" /rpt_type=DIRECT 12218. .12363
repeat_region	/partial /note="B1-MM-repeat" /rpt_family="SINE/Alu" /rpt_type=INVERTED 12799. .12954 /product="hexokinase II" /number=8 /used_in=y111666:hexIIrna /used_in=y111666:hexIICDS 12959. .13979 /number=8
intron	

```

repeat_region 13046..13083
/rpl_type=INVERTED
/rpl_unit=13046..13047
repeat_region 13873..13892
/rpl_type=INVERTED
/rpl_unit=13873..13874
exon 13980..14213
/product="hexokinase II"
/number=9
/intron 14214..15661
/usedin=Y11666:hexiImrna
/usedin=Y11666:hexiICDS
/number=9
exon 15662..15966
/product="hexokinase II"
/number=10
/intron 15967..17106
/usedin=Y11666:hexiICDS
/number=10
repeat_region 15988..16050
/rpl_type=INVERTED
/rpl_unit=15988..15989
exon 17107..17235
/product="hexokinase II"
/number=11
/intron 17256..17534
/usedin=Y11666:hexiImrna
/usedin=Y11666:hexiICDS
/number=11
exon 17535..17654
/product="hexokinase II"
/number=12
/intron 17655..20306
/usedin=Y11666:hexiImrna
/usedin=Y11666:hexiICDS
/number=12
repeat_region 17757..17897
/partial
/note="B1-MM-repeat"
/rpl_family="LINE/L1"
/rpl_type=INVERTED
20307..20402
/product="hexokinase II"
/number=13
/usedin=Y11666:hexiImrna
/usedin=Y11666:hexiICDS
/number=13
repeat_region 20608..20642
/rpl_type=INVERTED
/rpl_unit=20608..20609
20762..20930
/partial
/note="B2-repeat"
/rpl_family="SINE/B2"
/rpl_type=INVERTED
21356..21435
/product="hexokinase II"
/number=14
/usedin=Y11666:hexiImrna
/usedin=Y11666:hexiICDS
21456..21551
/number=14
exon 21552..21735
/product="hexokinase II"
/number=15
/usedin=Y11666:hexiImrna

```

```

Query Match 3.0%; Score 83; DB 10; Length 28646;
Best Local Similarity 100.0%; Pred No. 1.4e-32;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1360 GCTATGCTGACGGCGGTACCGCTGCTGCTGACCAACACCGGCGCCAGACACC 1419
Db 15756 GCTATGCTGACGGCGGTACCGCTGCTGCTGCTGACCAACACCGGCGCCAGACACC 15815
Qy 1420 CTGAGCTCTCTGAAGCTGAGCCA 1442
Db 15816 CTGAGCTCTCTGAAGCTGAGCCA 15838

```

```

RESULT 10
AC116811/C
LOCUS
DEFINITION Mus musculus clone RP24-356D15, LOW-PASS SEQUENCE SAMPLING.
AC116811.2
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, A.,
TITLE 1 (bases 1 to 67493)
JOURNAL Mus musculus, clone RP24-356D15
REFERENCE
AUTHORS 2 (bases 1 to 67493)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, A.,
Bouckhagalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeBellano, K., Dewar, K., Diaz, J., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gadyrna, S.,
Gilde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Melidris, N., Menus, L.,
Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollard, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Struss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Toppan, K., Traversman, A., Travis, N., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

```

```

TITLE
JOURNAL
REFERENCE
Submitted (02-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 67493)

```

Wed May 28 08:37:50 2003

us-09-808-743a-1.oligo.rge

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Camarato, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choquet, Y., Collangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goto, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamal, A., Karats, A., Kells, R., Lindblad-Toh, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Marquis, N.,
 Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N.,
 Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Melidim, J.,
 Menus, L., Mihova, T., Mienga, C. H., Murphy, T., Naylor, J., Nguyen, C.,
 Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Piere, N.,
 Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,
 Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S.,
 Schuppback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S.,
 Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J.,
 Vassiliev, H., Viel, R., Zembek, L., Zimmer, A., and Zody, M.
 Young, C., Zainoun, J., Zembek, L., Zimmer, A., and Zody, M.

TITLE

JOURNAL

Direct Submission
 Submitted (20-MAY-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 20, 2002 this sequence version replaced gi:19882065.

All repeats were identified using RepeatMasker:

smc, A.P.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WITB

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Project name: L25559

Center project name: 356_D-13

NOTE: This record contains 83 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely for
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 707: contig of 707 bp in length
 708 807: gap of 100 bp
 808 1524: contig of 717 bp in length
 1525 1624: gap of 100 bp
 1625 2360: contig of 736 bp in length
 2361 2460: gap of 100 bp
 2461 3170: contig of 710 bp in length
 3171 3270: gap of 100 bp
 3271 3990: contig of 720 bp in length
 3991 4090: gap of 100 bp
 4091 4804: contig of 714 bp in length
 4805 4904: gap of 100 bp
 4905 5580: contig of 676 bp in length
 5581 5680: gap of 100 bp
 5681 6413: contig of 733 bp in length
 6414 6513: gap of 100 bp
 6514 7211: contig of 698 bp in length
 7212 7311: gap of 100 bp
 7312 8039: contig of 728 bp in length
 8040 8139: gap of 100 bp
 8140 8852: contig of 713 bp in length
 8853 8952: gap of 100 bp
 8953 9685: contig of 733 bp in length
 9686 9785: gap of 100 bp

9786 10489: contig of 704 bp in length
 10490 10589: gap of 100 bp
 10590 11296: contig of 707 bp in length
 11297 11396: gap of 100 bp
 11397 12119: contig of 723 bp in length
 12120 12219: gap of 100 bp
 12220 12933: contig of 714 bp in length
 12934 13033: gap of 100 bp
 13034 13750: contig of 717 bp in length
 13751 13850: gap of 100 bp
 13851 14568: contig of 718 bp in length
 14569 14668: gap of 100 bp
 14669 15373: contig of 705 bp in length
 15374 15473: gap of 100 bp
 15474 16198: contig of 725 bp in length
 16199 16298: gap of 100 bp
 16299 17011: contig of 713 bp in length
 17012 17111: gap of 100 bp
 17112 17830: contig of 719 bp in length
 17831 17930: gap of 100 bp
 17931 18646: contig of 716 bp in length
 18647 18746: gap of 100 bp
 18747 19479: contig of 733 bp in length
 19480 19579: gap of 100 bp
 19580 20297: contig of 718 bp in length
 20298 20397: gap of 100 bp
 20398 21106: contig of 709 bp in length
 21107 21206: gap of 100 bp
 21207 21931: contig of 725 bp in length
 21932 22031: gap of 100 bp
 22032 22743: contig of 712 bp in length
 22744 22843: gap of 100 bp
 22844 23575: contig of 732 bp in length
 23576 23675: gap of 100 bp
 23676 24392: contig of 717 bp in length
 24393 24492: gap of 100 bp
 24493 25213: contig of 721 bp in length
 25214 25313: gap of 100 bp
 25314 26035: contig of 722 bp in length
 26036 26135: gap of 100 bp
 26136 26857: contig of 722 bp in length
 26858 26957: gap of 100 bp
 26958 27669: contig of 712 bp in length
 27670 27769: gap of 100 bp
 27770 28498: contig of 729 bp in length
 28499 28598: gap of 100 bp
 28599 29306: contig of 708 bp in length
 29307 29406: gap of 100 bp
 29407 30111: contig of 705 bp in length
 30112 30211: gap of 100 bp
 30212 30917: contig of 706 bp in length
 30918 31017: gap of 100 bp
 31018 31716: contig of 699 bp in length
 31717 31816: gap of 100 bp
 31817 32541: contig of 725 bp in length
 32542 32641: gap of 100 bp
 32642 33357: contig of 716 bp in length
 33358 33457: gap of 100 bp
 33458 34167: contig of 710 bp in length
 34168 34267: gap of 100 bp
 34268 34991: contig of 724 bp in length
 34992 35091: gap of 100 bp
 35092 35819: contig of 728 bp in length
 35820 35919: gap of 100 bp
 35920 36642: contig of 723 bp in length
 36643 36742: gap of 100 bp
 36743 37445: contig of 703 bp in length
 37446 37545: gap of 100 bp
 37546 38270: contig of 725 bp in length
 38271 38370: gap of 100 bp
 38371 39096: contig of 726 bp in length
 39097 39196: gap of 100 bp
 39197 39905: contig of 709 bp in length

Query Match 2.6%; Score 72; DB 10; Length 404;
 Best Local Similarity 100.0%; Pred. No. 2.1e-26;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 75 TCTTACCACATGGCTCTGCATGAGACCCCTCTGAGATTCTAGCGCTCCGAA 134
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 168 TCTTACCACATGGCTCTGCATGAGACCCCTCTGAGATTCTAGCGCTCCGAA 227
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Y 135 GCAGATGCAGAA 146
 ||||||||||||
 Db 228 GCAGATGCAGAA 239

RESULT 13
 RATHK2 1731 bp DNA linear ROD 07-NOV-2001
 LOCUS Rattus norvegicus HK2 gene for type II hexokinase, exon1 and
 DEFINITION Promoter region.
 ACCESSION D26393.1 GI:893403
 VERSION D26393.1 GI:893403
 KEYWORDS
 SOURCE Rattus norvegicus (strain:Wistar) DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 1731)
 Ichihara,J., Shinohara,Y., Kogure,K. and Terada,H.
 Nucleotide sequence of the 5'-flanking region of the rat type II
 hexokinase gene
 Biochim. Biophys. Acta 1260 (3), 365-368 (1995)

REFERENCE
 AUTHORS Ichihara,J., Shinohara,Y., Kogure,K. and Terada,H.
 TITLE Nucleotide sequence of the 5'-flanking region of the rat type II
 JOURNAL hexokinase gene
 MEDLINE Biochim. Biophys. Acta 1260 (3), 365-368 (1995)
 REFERENCE 2 (bases 1 to 1731)
 AUTHORS Shinohara,Y.
 TITLE Direct Submission
 JOURNAL Submitted (04-JAN-1994) Yasuo Shinohara, University of Tokushima,
 Faculty of Pharmaceutical Sciences, 1 Shomachi, Tokushima,
 Tokushima 770, Japan (E-mail:yasuoph.tokushima-u.ac.jp,
 Tel:81-886-33-7278, Fax:81-886-33-5196)
 FEATURES
 source
 1. 1731
 /organism="Rattus norvegicus"
 /strain="Wistar"
 /db_xref="taxon:10116"
 misc_signal
 56..61
 /note="CFI (transcription factor binding site)"
 misc_signal
 185..190
 /note="CFI (transcription factor binding site)"
 misc_signal
 464..469
 /note="PEA3 (transcription factor binding site)"
 misc_signal
 497..503
 /note="Pur (transcription factor binding site)"
 misc_signal
 844..856
 /note="AP2-PuF-AP2 (transcription factor binding site)"
 GC_signal
 938..943
 /note="AP2-PuF-AP2 (transcription factor binding site)"
 CAAT_signal
 1073..1077
 GC_signal
 1101..1106
 TATA_signal
 1129..1133
 exon
 1159..1683
 /number=1
 /evidence=experimental
 gene
 1621..1683
 /gene="HK2"
 CDS
 1621..>1683
 /gene="HK2"
 /EC_number="2.7.1.1"
 /codon_start=1
 /product="type II hexokinase"
 /protein_id="BA05409.1"
 /db_xref="GI:1526568"
 /translation="MASHMIACLFTELNQVQK"
 1696..1701
 CC_signal
 385 a 491 c 474 g 381 t
 BASE COUNT 385 a 491 c 474 g 381 t

ORIGIN
 Query Match 2.4%; Score 65; DB 10; Length 1731;
 Best Local Similarity 100.0%; Pred. No. 9.5e-23;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ATGATCGCCTCGCATATGATCGCTTATTTCAGGAGCTCAACCAACCAAGTCAG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1621 ATGATCGCCTCGCATATGATCGCTTATTTCAGGAGCTCAACCAACCAAGTCAG 1680
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Y 61 AAGGT 65
 |||||
 Db 1681 AAGGT 1685

RESULT 14
 AY082375 5067 bp DNA linear ROD 26-MAR-2002
 LOCUS Rattus norvegicus strain Sprague-Dawley type II hexokinase gene,
 DEFINITION exon 1 and partial cds.
 ACCESSION AY082375
 VERSION AY082375.1 GI:19743673
 KEYWORDS
 SOURCE Rattus norvegicus.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 5067)
 Rempel,A.
 Normal type II hexokinase promoter, first exon, and first intron
 from hepatocytes
 Unpublished
 2 (bases 1 to 5067)
 Rempel,A.
 Direct Submission
 Submitted (06-MAR-2002) Biological Chemistry, Johns Hopkins
 University, 725 N. Wolfe Street, Baltimore, MD 21205, USA
 FEATURES
 source
 1. 5067
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /cell_type="hepatocyte"
 mRNA
 4288..>4810
 /product="type II hexokinase"
 exon
 4288..4810
 /number=1
 4748..>4810
 /codon_start=1
 /product="type II hexokinase"
 /protein_id="AAL9251.1"
 /db_xref="GI:19743674"
 /translation="MASHMIACLFTELNQVQK"
 BASE COUNT 1228 a 1244 c 1254 g 1341 t
 ORIGIN

Query Match 2.4%; Score 65; DB 10; Length 5067;
 Best Local Similarity 100.0%; Pred. No. 8e-23;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ATGATCGCCTCGCATATGATCGCTTATTTCAGGAGCTCAACCAACCAAGTCAG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 4748 ATGATCGCCTCGCATATGATCGCTTATTTCAGGAGCTCAACCAACCAAGTCAG 4807
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Y 61 AAGGT 65
 |||||
 Db 4808 AAGGT 4812

RESULT 15
 RN019605 5150 bp DNA linear ROD 03-MAY-2001
 LOCUS RN019605

```

DEFINITION      Rattus norvegicus type II hexokinase gene, partial cds and promoter
                  region.
ACCESSION       U19605
VERSION         U19605.2
KEYWORDS        GI:13937421
SOURCE          Rattus norvegicus.
ORGANISM        Rattus norvegicus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                Rattus.
REFERENCE       1 (bases 1 to 5150)
AUTHORS         Mathupala, S.P., Kempel, A. and Pedersen, P.L.
TITLE           Glucose catabolism in cancer cells. Isolation, sequence, and
                activity of the promoter for type II hexokinase
JOURNAL         J. Biol. Chem. 270 (28), 16918-16925 (1995)
MEDLINE         95348123
PubMed         7622509
REFERENCE       2 (bases 1 to 5150)
AUTHORS         Mathupala, S.P.
TITLE           Direct Submission
JOURNAL         Submitted (09-JAN-1995) Biological Chemistry, Johns Hopkins
                University School of Medicine, 725 N. Wolfe Street, Baltimore, MD
                21205, USA
REFERENCE       3 (bases 1 to 5150)
AUTHORS         Lee, M.G., Ko, Y.H. and Pedersen, P.L.
TITLE           Direct Submission
JOURNAL         Submitted (03-MAY-2001) Biological Chemistry, Johns Hopkins
                University School of Medicine, 725 N. Wolfe Street, Baltimore, MD
                21205, USA
REMARK          Sequence update by submitter
COMMENT          On May 3, 2001 this sequence version replaced gi:901870.
FEATURES        source
                  1..5150
                    /organism="Rattus norvegicus"
                    /strain="Sorague-Dawley"
                    /db_xref="taxon:10116"
                    /clone="29-1/Xbat"
                    /sex="female"
                    /cell_line="AS-30D"
                    /tissue_type="hepatoma, ascites"
                    /clone_id="Lambda-PhiXII/AS-30D genomic"
                    4284..4288
                    4340..4344
                    4370..4893
                    /number=1
                    4831..>4893
                    /codon_start=1
                    /product="type II hexokinase"
                    /protein_id="AB09025.1"
                    /db_xref="GI:1568599"
                    /translation="MRASHMIACTFELNQNVOK"
                    4894..>5150
                    /number=1
Intron
BASE COUNT     1249 a 1263 c 1277 g 1361 t
ORIGIN
Query Match    2.4%; Score 65; DB 10; Length 5150;
Best Local Similarity 100.0%; Pred. No. 7.9e-23;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGATGCCCTGCATATGATGCGCTGCTATTTCACGAGCTCAACCAAAACCAAGTGCAG 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4831 ATGATGCCCTGCATATGATGCGCTGCTATTTCACGAGCTCAACCAAAACCAAGTGCAG 60
QY 61 AAGGT 65
    |||||
Db 4891 AAGGT 4895

```

Search completed: May 27, 2003, 16:34:47
 Job time : 8826 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2003, 11:57:02 ; Search time 596 Seconds
(without alignments)
10406.046 Million cell updates/sec

Title: US-09-808-743a-1

Perfect score: 2754
Sequence: 1 atgacgcctccatcatgat.....gggagcgtgacagagatag 2754

Scoring table: OLIGO NMC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 50

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

N_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2754	100.0	2754	AA167594	Rat mutant type II
2	2753	100.0	2770	AA178599	AS-30D tumour type
3	2550	92.6	3635	AAV00125	Rat hexokinase II
4	2550	92.6	3635	AAV00086	Rat hexokinase II
5	102	3.7	299	AA180522	AS-30D tumour type
6	65	2.4	5150	AA178598	AS-30D tumour type

ALIGNMENTS

RESULT 1
AA167594
ID AA167594 standard; DNA: 2754 BP.
XX
AC AA167594;
XX
DT 27-FEB-2002 (first entry)
XX
DE Rat mutant type II hexokinase encoding DNA.
XX
KW Hexokinase: cell proliferation; glycolytic tumour; cancer; mutant;
type II hexokinase; cytosolic; glycolysis inhibitor; gene therapy; ds.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT CDS 1..2754
FT /tag= a
FT /product= "mutant type II hexokinase"
XX
PN WO20016667-A1.
XX
PD 20-SEP-2001.
XX
PF 14-MAR-2001; 2001WO-US08335.
XX
PR 14-MAR-2000; 2000US-189222P.
XX
PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Pedersen PL, Mathupala SP;
XX
DR WPI: 2002-049005/06.
DR P-PSDB: AAG66000.
XX
PT Inhibiting proliferation of highly glycolytic tumors, e.g. in gastric
cancer, hepatoma, colorectal cancer or lung cancer, by contacting cells
with antisense molecules that hybridize with a nucleic acid encoding
hexokinase -
XX
PS Claim 2: Fig 6; 59pp; English.
XX
CC The invention provides a method for inhibiting proliferation of tumour
cells characterized by having a highly glycolytic phenotype. The method
involves contacting the cells with an antisense polynucleotide or
oligonucleotide that hybridizes with a mRNA encoding a hexokinase under
conditions that allow hybridization of the antisense polynucleotide with
the mRNA, thus inhibiting the proliferation of tumour cells. The method
is useful for inhibiting proliferation of highly glycolytic tumors or
for modulating the expression of a hexokinase in highly glycolytic
tumours. In particular, the cellular proliferative disorder comprises low
grade astrocytoma, anaplastic astrocytoma, glioblastoma, medulloblastoma,
gastric cancer, hepatoma, colorectal cancer, colorectal adenoma, acute
myelogenous leukemia, lung cancer, renal cancer, leukemia, breast cancer,
prostate cancer, endometrial cancer, bone cancer, squamous cell cancer,
and neuroblastoma. The present sequence represents the DNA encoding a rat
mutant type II hexokinase.
XX
SQ Sequence 2754 BP; 642 A; 659 C; 869 G; 584 T; 0 other;

Query Match 100.0%; Score 2754; DB 24; Length 2754;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2754; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGATGCGCTCCCATATGATGCTGCTATTCACGAGAGCTCAACCAAAACCAAGTCAG 60
|||||
DB 1 ATGATGCGCTCCCATATGATGCTGCTATTCACGAGAGCTCAACCAAAACCAAGTCAG 60
|||||
OY 61 AAGGTTGACCAATTTCTCTACACATGCGCTCTCTAGATGAGACCCCTTCTGAGATTTCCT 120
|||||
DB 61 AAGGTTGACCAATTTCTCTACACATGCGCTCTCTAGATGAGACCCCTTCTGAGATTTCCT 120
|||||

OY 121 AGGGGTTTCGGAGAGATGGAGAAAGGGCTTAGAGCTTACACGACCTTACAGCACT 180
DB 121 AGGGGGTTCCGGAGAGATGGAGAAAGGGCTTAGAGCTTACACGACCTTACAGCACT 180
OY 181 GTGAAATGTTGGCTTACCTTGTGTAGGTCAACTCCGATGGGAGACAGAACATGGGAGTTTC 240
DB 181 GTGAAATGTTGGCTTACCTTGTGTAGGTCAACTCCGATGGGAGACAGAACATGGGAGTTTC 240
OY 241 CTGGCTCTGATCTTTGGAGGAAACCACTTCGCTGTCTCCAGTAAAGGTGACGGACAAT 300
DB 241 CTGGCTCTGATCTTTGGAGGAAACCACTTCGCTGTCTCCAGTAAAGGTGACGGACAAT 300
OY 301 GGCCCTCCAGAGATGGAGATGGAGAACAGATCTACGCCATCTCTTGAGACATCATGTGGG 360
DB 301 GGCCCTCCAGAGATGGAGATGGAGAACAGATCTACGCCATCTCTTGAGACATCATGTGGG 360
OY 361 GGCAGTGGAAACCCAGCTGTTTGACCAATCCGCCAATCCCTGGCCAACTTCATGGACAA 420
DB 361 GGCAGTGGAAACCCAGCTGTTTGACCAATCCGCCAATCCCTGGCCAACTTCATGGACAA 420
OY 421 CTACAAATCAAGAGAGAAAGCTCCCTGGGTTTCACTTCTGTGTCCTCCGACCAAG 480
DB 421 CTACAAATCAAGAGAGAAAGCTCCCTGGGTTTCACTTCTGTGTCCTCCGACCAAG 480
OY 481 ACAAACCTGATGAGAGTCTTTTGTCTCTGTGACTAAGGGGTTCAAGTCCAGTGGCGTG 540
DB 481 ACAAACCTGATGAGAGTCTTTTGTCTCTGTGACTAAGGGGTTCAAGTCCAGTGGCGTG 540
OY 541 GAAGCACAAGATGTGTGGAGCTGATCCGGAAAGTTATCCAGCCGAGAGGGGACTTTGAC 600
DB 541 GAAGCACAAGATGTGTGGAGCTGATCCGGAAAGTTATCCAGCCGAGAGGGGACTTTGAC 600
OY 601 ATTGACATTTGGGCGGTGTGATGACACAGTGGGACCATGATGATGTTGGCTATGAT 660
DB 601 ATTGACATTTGGGCGGTGTGATGACACAGTGGGACCATGATGATGTTGGCTATGAT 660
OY 661 GATCAAACTCGGAGATTTGTCTCATTTGGGCACTGGCAGCAACGCTGCTACATGGAG 720
DB 661 GATCAAACTCGGAGATTTGTCTCATTTGGGCACTGGCAGCAACGCTGCTACATGGAG 720
OY 721 GAAATGCTCATATGATGATGGTGGAGAGATGAGGGGCGCATGTGCATCAATGAGG 780
DB 721 GAAATGCTCATATGATGATGGTGGAGAGATGAGGGGCGCATGTGCATCAATGAGG 780
OY 781 TGGGGAGCCTTTGGGGAGCAGCGGTACACTCAATGACATCCGAACCGAGTTTGGACCGAG 840
DB 781 TGGGGAGCCTTTGGGGAGCAGCGGTACACTCAATGACATCCGAACCGAGTTTGGACCGAG 840
OY 841 ATGCAATGGGCTCGCTGAAACCTTGGGAAAGCAGCTGTTTGAAGAATGATTTAGCGGGATG 900
DB 841 ATGCAATGGGCTCGCTGAAACCTTGGGAAAGCAGCTGTTTGAAGAATGATTTAGCGGGATG 900
OY 901 TGCATGGGCGAGCTGTGCTGAGCTCATCTGCTGAGAGATGGCCAAAGCAGAGCTGTTGTC 960
DB 901 TGCATGGGCGAGCTGTGCTGAGCTCATCTGCTGAGAGATGGCCAAAGCAGAGCTGTTGTC 960
OY 961 CAAGGAAACTAGCCACAGAACTCTTACCACTGCTCCTTTCGAGACCAAGATGTTCTG 1020
DB 961 CAAGGAAACTAGCCACAGAACTCTTACCACTGCTCCTTTCGAGACCAAGATGTTCTG 1020
OY 1021 GATATTGGAAGAGATTAAGATGGAATCGAAGAGGCTTACCAAAATCCTGTGGCGGT 1080
DB 1021 GATATTGGAAGAGATTAAGATGGAATCGAAGAGGCTTACCAAAATCCTGTGGCGGT 1080
OY 1081 CTGAATTCATTCAGAGAGATTTGTGTGCGACAGCAGCAAGATCTGCCAGATTTGTGTCAG 1140
DB 1081 CTGAATTCATTCAGAGAGATTTGTGTGCGACAGCAGCAAGATCTGCCAGATTTGTGTCAG 1140
OY 1141 CGCTGGCGCAGTCTGTGCGCAGCCACCTGGCGGGGTGCTGTGGGGAATCAAGAGAAC 1200
DB 1141 CGCTGGCGCAGTCTGTGCGCAGCCACCTGGCGGGGTGCTGTGGGGAATCAAGAGAAC 1200

OY 1201 AAGGCGAGAGAGCAGCTTCGCTCCACCATCGGTGCTCATGAGCTCCGTTACAGAAACAT 1260
DB 1201 AAGGCGAGAGAGCAGCTTCGCTCCACCATCGGTGCTCATGAGCTCCGTTACAGAAACAT 1260
OY 1261 CCCCATTTCGCAAGCGCTCTCATTAAGCACTGAGAGAGCTGTGTGCGGACTGTGATGTC 1320
DB 1261 CCCCATTTCGCAAGCGCTCTCATTAAGCACTGAGAGAGCTGTGTGCGGACTGTGATGTC 1320
OY 1321 CGGTTCCCTCGCTGAGAGATGGCAGCGGCAAGGGGGCTGTATAGTGTGACGGCGGTGCT 1380
DB 1321 CGGTTCCCTCGCTGAGAGATGGCAGCGGCAAGGGGGCTGTATAGTGTGACGGCGGTGCT 1380
OY 1381 TACCGTGTGGCTGACCAACACCGGGCCCGCAGAGAACCTTGAGTCTTGAAGCTGAGC 1440
DB 1381 TACCGTGTGGCTGACCAACACCGGGCCCGCAGAGAACCTTGAGTCTTGAAGCTGAGC 1440
OY 1441 CACGAGCAGCTTCGAGAGATTAAGAGAAATGAAGGTGGAATGAGAGCAGGTCTGAGC 1500
DB 1441 CACGAGCAGCTTCGAGAGATTAAGAGAAATGAAGGTGGAATGAGAGCAGGTCTGAGC 1500
OY 1501 AAGGAGACGATGGCGTCCGCTGTGAAGATGCTGCCACTTACGTGTGACCACTCA 1560
DB 1501 AAGGAGACGATGGCGTCCGCTGTGAAGATGCTGCCACTTACGTGTGACCACTCA 1560
OY 1561 GATGCAACAGAGAAAGAGACTTCTGGCCTTGGATCTTGGAGACAAACATTCGGGTC 1620
DB 1561 GATGCAACAGAGAAAGAGACTTCTGGCCTTGGATCTTGGAGACAAACATTCGGGTC 1620
OY 1621 CTGCTGTGCTGTGCTGATATGAGAGAGGCGCTGAGAGATGATTAAGAGATCTTAC 1680
DB 1621 CTGCTGTGCTGTGCTGATATGAGAGAGGCGCTGAGAGATGATTAAGAGATCTTAC 1680
OY 1681 TCCATCCACAGAGAGATTAAGAGAGCACTGGGGAAGAGCTTCTGACCACTTGTCCAG 1740
DB 1681 TCCATCCACAGAGAGATTAAGAGAGCACTGGGGAAGAGCTTCTGACCACTTGTCCAG 1740
OY 1741 TGCATTTGGGCACTTCCTGAGAGATGAGGCAATGAGGCGGTGCTCCGTTGGGTTTC 1800
DB 1741 TGCATTTGGGCACTTCCTGAGAGATGAGGCAATGAGGCGGTGCTCCGTTGGGTTTC 1800
OY 1801 ACATTCCTCCTTCCTTGGCAGAGAAACAGCCTTAGACCAAGACATCCTCAAGTGACA 1860
DB 1801 ACATTCCTCCTTCCTTGGCAGAGAAACAGCCTTAGACCAAGACATCCTCAAGTGACA 1860
OY 1861 AAGGATTCAGAGCACTTGTGCTGCGAGGGGTGAGAGATGTGTGCTTCTCTGAAGAACG 1920
DB 1861 AAGGATTCAGAGCACTTGTGCTGCGAGGGGTGAGAGATGTGTGCTTCTCTGAAGAACG 1920
OY 1921 ATTACCGGCGAGAGAGATTTGACCTGATGTGTGCGGTGTGAATGACACATTTGG 1980
DB 1921 ATTACCGGCGAGAGAGATTTGACCTGATGTGTGCGGTGTGAATGACACATTTGG 1980
OY 1981 ACTATGATGACTTGTGCTACAGAAACCTCCTCACTGTGAAGTGGCTCATTTGTTGGACC 2040
DB 1981 ACTATGATGACTTGTGCTACAGAAACCTCCTCACTGTGAAGTGGCTCATTTGTTGGACC 2040
OY 2041 GGAACCAACGCTGCTTAACATGGAAGAGATGCTTAATGTGAGCTGTGAGGAGAGAG 2100
DB 2041 GGAACCAACGCTGCTTAACATGGAAGAGATGCTTAATGTGAGCTGTGAGGAGAGAG 2100
OY 2101 GGAGGATGTGTCAACATGAGAGTGGGAGCACTTTGGGGAACATGCTGCTGATGAT 2160
DB 2101 GGAGGATGTGTGTCAACATGAGAGTGGGAGCACTTTGGGGAACATGCTGCTGATGAT 2160
OY 2161 TTGCGAGCCGTTTGTATGTTGCTGTGATGAGCTTTCTCTCAACCTTGGCAACAGAG 2220
DB 2161 TTGCGAGCCGTTTGTATGTTGCTGTGATGAGCTTTCTCTCAACCTTGGCAACAGAG 2220
OY 2221 TTGCAAGAGATATCAGGGGCACTTGTGAGAGATTTGGCGCAACATTTCTCAT 2280
DB 2221 TTGCAAGAGATATCAGGGGCACTTGTGAGAGATTTGGCGCAACATTTCTCAT 2280
OY 2281 TTCACGAAGCGGGGCTGCTTTCGAGGCGCATCTCAGAGCGGCTCAAGCAAGAGGA 2340


```

Db 2281 TTACGAGAGCGGGGCTGCTTCGAGGCGCATCTGAGAGCGGCTCAAGACAAAGGGA 2340
OY 2341 ATCTGAAACTAAGTCTCTGCTCAGATAGAGAGCGAGCTGCTAGGCTGTACAGGTT 2400
Db 2341 ATCTGAAACTAAGTCTCTGCTCAGATAGAGAGCGAGCTGCTAGGCTGTACAGGTT 2400
OY 2401 CGTGGCATCTCTGCGCCACCTAGAGGGGTGAGAGACAGTGCATGACAGCATCATGTAAG 2460
Db 2401 CGTGGCATCTCTGCGCCACCTAGAGGGGTGAGAGACAGTGCATGACAGCATCATGTAAG 2460
OY 2461 GAGGTGTGACAGTGTGCTGCGCGCGCGCTGACAGCTGTGTGGCGAGCATGCGCGC 2520
Db 2461 GAGGTGTGACAGTGTGCTGCGCGCGCGCTGACAGCTGTGTGGCGAGCATGCGCGC 2520
OY 2521 GTAGTGGCAAGATAGAGAGAAACCGTGGGTGAGACACCCCAAGTACATGAGCGCTG 2580
Db 2521 GTAGTGGCAAGATAGAGAGAAACCGTGGGTGAGACACCCCAAGTACATGAGCGCTG 2580
OY 2581 GACGGGACTCTGTATAGCTTCACTTCTCACTTGGCCAAAGTCAATGACAGCGTGA 2640
Db 2581 GACGGGACTCTGTATAGCTTCACTTCTCACTTGGCCAAAGTCAATGACAGCGTGA 2640
OY 2641 GATCTGCTCCGAATGTAGCTGCTCTCTGGAATCCGAGAGCGAGCGAGTGGAGGA 2700
Db 2641 GATCTGCTCCGAATGTAGCTGCTCTCTGGAATCCGAGAGCGAGCGAGTGGAGGA 2700
OY 2701 GCAGCTCTCATCTGCGCTGCGCGCATCCGCGAGCGTGGGAGAGATAG 2754
Db 2701 GCAGCTCTCATCTGCGCTGCGCGCATCCGCGAGCGTGGGAGAGATAG 2754

RESULT 2
AAT78599
ID AAT78599 standard; cDNA; 2770 BP.
AC AAT78599;
XX 31-OCT-1997 (first entry)
XX AS-30D tumour Type II hexokinase encoding cDNA.
XX Response element; 2-DNA; neoplasia; hexokinase II; glycolysis;
XX cancer; gene therapy; diabetes; tumour; rat; ss.
OS Rattus rattus.
FH key
FT CDS Location/Qualifiers
FT 18..2770
FT /*tag=
FT /product= AS-30D_Type II hexokinase
FT /note= "The last base of the stop codon is not shown"
PN MO9704104-A2.
XX
XX PD 06-FEB-1997.
XX 12-JUL-1996; 96MO-US11673.
XX 14-JUL-1995; 95US-0001199.
XX (UYJO ) UNIV JOHNS HOPKINS.
XX PI Mathupala SP, Pedersen PL, Rempel A.
XX WPI: 1997-132643/12.
XX P-PSDB; AAM23793.
XX
XX New transcription regulating fragments of hexokinase II DNA conty.
XX response element - and methods for diagnosis or treatment of
XX neoplasias that over-express hexokinase II and for regulating
XX glycolysis

```

```

PS Disclosure: Fig 6; 104pp; English.
XX
CC The present sequence represents the AS-30D tumour Type II hexokinase
CC gene, which is a new isolated hexokinase II. DNA fragments from the
CC promoter region (shown in AAT78599), coding sequences and introns are
CC able to regulate transcription of a downstream open reading frame if
CC they contain at least one response element (transcription factor
CC binding site). A method has been produced for diagnosing a neoplasia
CC that over-expresses hexokinase by detecting a copy number greater than
CC 2 for the hexokinase II gene. The DNA fragments may be coupled to a
CC reporter gene and used to screen for potential drugs that affect
CC regulated transcription of tumour hexokinase II. Alternatively they may
CC be coupled to a toxic gene and used to treat cells that over-express
CC hexokinase II, such as those present in patients with cancer. They may
CC also be used in gene therapy to treat diabetes. The DNA fragments can
CC increase glycolysis in cells and express homologous or heterologous
CC protein. Probes of the DNA fragment are used in the method for
CC diagnosing a neoplasia that over-expresses hexokinase. The new response
CC elements are active only in tumours, not in normal cells.
XX
SQ Sequence 2770 BP; 645 A; 666 C; 874 G; 585 T; 0 other;
Query Match 100.0%; Score 2753; DB 18; Length 2770;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGATCCCTCGCATATGATCGCTCTTATTCAGGAGCTTCACCAAAACCAAGTCAG 60
Db 18 ATGATCCCTCGCATATGATCGCTCTTATTCAGGAGCTTCACCAAAACCAAGTCAG 60
OY 61 AAGGTGACCAATTTCTTACCAATGGCTCTCTCAGATGAGACCTTGTGAGATTTCT 77
Db 78 AAGGTGACCAATTTCTTACCAATGGCTCTCTCAGATGAGACCTTGTGAGATTTCT 77
OY 121 AGCGGTTCCGGAAGAGATGAGAAAGGGCTAGAGCTCCAGACCTTACAGAGCT 137
Db 121 AGCGGTTCCGGAAGAGATGAGAAAGGGCTAGAGCTCCAGACCTTACAGAGCT 137
OY 138 AGCGGTTCCGGAAGAGATGAGAAAGGGCTAGAGCTCCAGACCTTACAGAGCT 180
Db 138 AGCGGTTCCGGAAGAGATGAGAAAGGGCTAGAGCTCCAGACCTTACAGAGCT 180
OY 181 GTGAAATGTTGCCACCTTGTGAGGTCAACTCCGATGAGAGACATGAGAGTTC 197
Db 181 GTGAAATGTTGCCACCTTGTGAGGTCAACTCCGATGAGAGACATGAGAGTTC 197
OY 198 GTGAAATGTTGCCACCTTGTGAGGTCAACTCCGATGAGAGACATGAGAGTTC 240
Db 198 GTGAAATGTTGCCACCTTGTGAGGTCAACTCCGATGAGAGACATGAGAGTTC 240
OY 241 CTGGCTGTGATCTTGGAGAGACCAACTCCGATGAGAGTTCGAGAGTTC 257
Db 241 CTGGCTGTGATCTTGGAGAGACCAACTCCGATGAGAGTTCGAGAGTTC 257
OY 258 CTGGCTGTGATCTTGGAGAGACCAACTCCGATGAGAGTTCGAGAGTTC 300
Db 258 CTGGCTGTGATCTTGGAGAGACCAACTCCGATGAGAGTTCGAGAGTTC 300
OY 301 GGCTCCAGAGATGAGATGAGAGACCAACTCCGATGAGAGTTCGAGAGTTC 317
Db 301 GGCTCCAGAGATGAGATGAGAGACCAACTCCGATGAGAGTTCGAGAGTTC 317
OY 318 GGCTCCAGAGATGAGATGAGAGACCAACTCCGATGAGAGTTCGAGAGTTC 360
Db 318 GGCTCCAGAGATGAGATGAGAGACCAACTCCGATGAGAGTTCGAGAGTTC 360
OY 361 GGCAGTGAACCCAGCTGTTTGAACACATCGCCGATGAGAGTTCGAGAGTTC 377
Db 361 GGCAGTGAACCCAGCTGTTTGAACACATCGCCGATGAGAGTTCGAGAGTTC 377
OY 378 GGCAGTGAACCCAGCTGTTTGAACACATCGCCGATGAGAGTTCGAGAGTTC 420
Db 378 GGCAGTGAACCCAGCTGTTTGAACACATCGCCGATGAGAGTTCGAGAGTTC 420
OY 421 CTCAATATCAAGAGAGAGCTCCCTGAGTTCACCTTCTGCTCCCTCCACAG 437
Db 421 CTCAATATCAAGAGAGAGCTCCCTGAGTTCACCTTCTGCTCCCTCCACAG 437
OY 438 CTCAATATCAAGAGAGAGCTCCCTGAGTTCACCTTCTGCTCCCTCCACAG 480
Db 438 CTCAATATCAAGAGAGAGCTCCCTGAGTTCACCTTCTGCTCCCTCCACAG 480
OY 481 ACAAACTGAGATGAGATGTTTGTCTGCTGAGACTAAGGGTTCAAGTCCAGTGC 497
Db 481 ACAAACTGAGATGAGATGTTTGTCTGCTGAGACTAAGGGTTCAAGTCCAGTGC 497
OY 498 ACAAACTGAGATGAGATGTTTGTCTGCTGAGACTAAGGGTTCAAGTCCAGTGC 540
Db 498 ACAAACTGAGATGAGATGTTTGTCTGCTGAGACTAAGGGTTCAAGTCCAGTGC 540
OY 541 GAAGGCAAGATGAGTGTGACCTGATCCGAGAGTTCATCCAGCGCAGAGGACTTTC 557
Db 541 GAAGGCAAGATGAGTGTGACCTGATCCGAGAGTTCATCCAGCGCAGAGGACTTTC 557
OY 558 GAAGGCAAGATGAGTGTGACCTGATCCGAGAGTTCATCCAGCGCAGAGGACTTTC 600
Db 558 GAAGGCAAGATGAGTGTGACCTGATCCGAGAGTTCATCCAGCGCAGAGGACTTTC 600
OY 601 ATTGACATGTTGGCCGCTGATGAGTTCAGAGTTCGAGAGTTCGAGAGTTC 617
Db 601 ATTGACATGTTGGCCGCTGATGAGTTCAGAGTTCGAGAGTTCGAGAGTTC 617
OY 618 ATTGACATGTTGGCCGCTGATGAGTTCAGAGTTCGAGAGTTCGAGAGTTC 660
Db 618 ATTGACATGTTGGCCGCTGATGAGTTCAGAGTTCGAGAGTTCGAGAGTTC 660
OY 661 GATCAGAACTGCGAGATGTTGCTCATTTGGGCACTGCGAGAGCTGATGAG 677
Db 661 GATCAGAACTGCGAGATGTTGCTCATTTGGGCACTGCGAGAGCTGATGAG 677
OY 678 GATCAGAACTGCGAGATGTTGCTCATTTGGGCACTGCGAGAGCTGATGAG 720
Db 678 GATCAGAACTGCGAGATGTTGCTCATTTGGGCACTGCGAGAGCTGATGAG 720

```

QY	1801	ACATTCCTCTCCCTTCCAGAGAAACACCTAGACACGACCATCTCTCTCAAGTGGACA	1860
Db	1818	ACATTCCTCTCCCTTCCAGAGAAACACCTAGACACGACCATCTCTCTCAAGTGGACA	1877
QY	1861	AAGGATTCAGAGGATCTGGCTCCGAGGGTGAACATGTGATCCTTCTCTAAGAAAGCG	1920
Db	1878	AAGGATTCAGAGGATCTGGCTCCGAGGGTGAACATGTGATCCTTCTCTAAGAAAGCG	1937
QY	1921	ATTCAACGGGGAGAGAGATTGACCTTGATGTGTTCCCTGTGAATGACACAGTTGGG	1980
Db	1938	ATTCAACGGGGAGAGAGATTGACCTTGATGTGTTCCCTGTGAATGACACAGTTGGG	1997
QY	1981	ACTATGATGACTTGTGGCTAGAGAGACCCATCAGTGTGAATGTGGCTCATTTGTGGACCC	2040
Db	1998	ACTATGATGACTTGTGGCTAGAGAGACCCATCAGTGTGAATGTGGCTCATTTGTGGACCC	2057
QY	2041	GGAGCAAGCGCTGTACATGAGAGATGCGTAAATGTGGAGCTGTGTGACGAGAGAGAG	2100
Db	2058	GGAGCAAGCGCTGTACATGAGAGATGCGTAAATGTGGAGCTGTGTGACGAGAGAGAG	2117
QY	2101	GGAGCAATGTGTCTCAACTATGAGAGTGGAGAGACATTTGGGGACATATGGCTCCCTGATGAC	2160
Db	2118	GGAGCAATGTGTCTCAACTATGAGAGTGGAGAGACATTTGGGGACATATGGCTCCCTGATGAC	2177
QY	2161	TTGGCGACCGGTGTTGATGTGTGCTGTGAGATGAGCTTTTCTCTCAACCCCTGCAACAGAGG	2220
Db	2178	TTGGCGACCGGTGTTGATGTGTGCTGTGAGATGAGCTTTTCTCTCAACCCCTGCAACAGAGG	2237
QY	2221	TTTCGAGAGATGATCAGACGGGCATGTACTTGGAGAGATTTGTCCGCAACATTTCTCATGAT	2280
Db	2238	TTTCGAGAGATGATCAGACGGGCATGTACTTGGAGAGATTTGTCCGCAACATTTCTCATGAT	2297
QY	2281	TTTCACGAGACGGGGGCTGTCTTTTCCGAGGCGGCATGCTCAAGCGCCTCANGACAAAGGGGA	2340
Db	2298	TTTCACGAGACGGGGGCTGTCTTTTCCGAGGCGGCATGCTCAAGCGCCTCANGACAAAGGGGA	2357
QY	2341	ATCTCTGAACACTAAGTCTCTGTCTCTCAATGATAGAGAGGACATGCTTAAGCCTGTACAGGTT	2400
Db	2358	ATCTCTGAACACTAAGTCTCTGTCTCTCAATGATAGAGAGGACATGCTTAAGCCTGTACAGGTT	2417
QY	2401	CGTGGCATCTCGCGCCACCTAGAGGCTGTGAGAGACACGTGCGATGACACAGCATCATCTGGAGAG	2460
Db	2418	CGTGGCATCTCGCGCCACCTAGAGGCTGTGAGAGACACGTGCGATGACACAGCATCATCTGGAGAG	2477
QY	2461	GAGGTGTGCACTGTGGTTGCCGGGGCGGTGTGCACAGCTCTGTGGCGCAAGCATATGGCGGCC	2520
Db	2478	GAGGTGTGCACTGTGGTTGCCGGGGCGGTGTGCACAGCTCTGTGGCGCAAGCATATGGCGGCC	2537
QY	2521	GTAGTGGACAGATTAAGAGAAACCTGTGGGCTGTGACCAACCCCAAGTATGACAGTGGCGTG	2597
Db	2538	GTAGTGGACAGATTAAGAGAAACCTGTGGGCTGTGACCAACCCCAAGTATGACAGTGGCGTG	2604
QY	2581	GACGGGACTCTGTATTAAAGCTTCATCTCTCACTTGTGCCAAGGTATGATGATGAGACGGTGA	2640
Db	2598	GACGGGACTCTGTATTAAAGCTTCATCTCTCACTTGTGCCAAGGTATGATGATGAGACGGTGA	2657
QY	2641	GATCTGGCTCGAATATGAGAGAGTGTCTCTCTGAAATCCGAGAGCGGACAGTGGAAAGGA	2700
Db	2658	GATCTGGCTCGAATATGAGAGAGTGTCTCTCTGAAATCCGAGAGCGGACAGTGGAAAGGA	2717
QY	2701	GCAGCTCTATCATCTGCGGTGGCTGTGCGCATCCGGAGAGGCTGGCGACAGAGATA	2753
Db	2718	GCAGCTCTATCATCTGCGGTGGCTGTGCGCATCCGGAGAGGCTGGCGACAGAGATA	2770
RESULT 3			
AAV00125			
AAV00125 standard; cDNA: 3635 BP.			
AAV00125:			
17-MAR-1998 (first entry)			

DE Rat hexokinase II encoding cDNA.
XX Hexokinase; competitive enzyme; inhibitor; insulin; diabetes; B cell;
KM glucose; glucokinase; ss.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT CDS 198..2951
FT /*lag= a
FT /product= Hexokinase_II
XX
PN W09726322-A2.
XX
XX 24-JUL-1997.
XX
XX 17-JAN-1997; 97WO-US00786.
XX
XX 19-JAN-1996; 96US-0588976.
XX
XX (BETA-) BETAGENE INC.
XX (UMMS) UNIV MICHIGAN STATE.
XX (TEXA) UNIV TEXAS SYSTEM.
XX
XX Becker TC, Han H, Newgard CB, Wilson JE;
XX
XX WPI: 1997-365327/35.
XX P-PSDB: AAW37437.
XX
XX Mammalian cell with reduced activity of low Km hexokinase - caused
PT by competitive enzyme inhibitor, particularly an insulin secreting
PT cell for in vitro or in vivo insulin production
XX
XX Claim 14; Page 186-189; 254pp; English.
XX
XX A mammalian cell has been developed which includes enough of a
CC competitive inhibitor of low Km (Michaelis constant) hexokinase to
CC reduce the hexokinase activity in the cell. The present sequence
CC encodes rat hexokinase II. The cells are particularly used to produce
CC insulin, in response to glucose or other secretagogues, either in vitro
CC or in vivo (for treating diabetes), but may also be used to produce
CC many other therapeutic proteins, e.g. from a gene under control of the
CC insulin promoter and thus responsive to glucose. Inhibition of the
CC inhibitor can be used more generally to reduce the growth rate of the
CC cells, e.g. hybridomas that produce antibodies.
XX
XX Sequence 3635 BP; 816 A; 875 C; 1098 G; 846 T; 0 other;
SQ
Query Match 92.6%; Score 2550; DB 18; Length 3635;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2750; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGATCGCCTCGCATATGATGCTGCTTATTACAGGAGCTAACCAAAACCAAGTCAG 60
DB 198 ATGATCGCCTCGCATATGATGCTGCTTATTACAGGAGCTAACCAAAACCAAGTCAG 257
QY 61 AAGGTGACCAATTTCTACACACATGCTCTCAGATGAGACCTTCTGAGATTTCT 120
DB 258 AAGGTGACCAATTTCTACACACATGCTCTCAGATGAGACCTTCTGAGATTTCT 317
QY 121 AAGCGGTTCCGAGAGAGATGAGAAAGGCTAGAGAGCTACACGACCCCTACAGAGCT 180
DB 318 AAGCGGTTCCGAGAGAGATGAGAAAGGCTAGAGAGCTACACGACCCCTACAGAGCT 377
QY 181 GTGAAAATGTTCCTTCTTGTGAGTCACTCCGGATGGGACAGAAATGGGAGTTTC 240
DB 378 GTGAAAATGTTCCTTCTTGTGAGTCACTCCGGATGGGACAGAAATGGGAGTTTC 437
QY 241 CTGGCTCTGATCTTGGAGAGACCAACTTCGCTGCTCCGAGTAAAGGTGACGAGCAAT 300
DB 438 CTGGCTCTGATCTTGGAGAGACCAACTTCGCTGCTCCGAGTAAAGGTGACGAGCAAT 497
QY 301 GGCCTTCAGAGAGATGAGAGAAACGAGATCTAGCCATCTTGAGGACATCATGCGG 360

|||||
DB 498 GGCCTTCAGAGAGATGAGAGAAACCAACCAATCTACGCCATCCCTGAGGACATATGCGG 557
QY 361 GGCAGTGAACCCAGCTGTTTGTACACACATGCGCAATGCTGGCCAACTTCATGACAAG 420
DB 558 GGCAGTGAACCCAGCTGTTTGTACACACATGCGCAATGCTGGCCAACTTCATGACAAG 617
QY 421 CTACAAATCAAGAGAAAGAGCTCCCTGCTGCTTCCCTTCCTGCTCCCTGACACAG 480
DB 618 CTACAAATCAAGAGAAAGAGCTCCCTGCTGCTTCCCTTCCTGCTCCCTGACACAG 677
QY 481 ACAAACTGATGAGAGATTTTGTCTGCTGAGACTAAGGGGTTCAAGTCCAGTGGGCTG 540
DB 678 ACAAACTGATGAGAGATTTTGTCTGCTGAGACTAAGGGGTTCAAGTCCAGTGGGCTG 737
QY 541 GAAGGACAGATGTGTGAGACCTGATCCGAGAGTTATCCAGCCAGAGGGGACTTTGAC 600
DB 738 GAAGGACAGATGTGTGAGACCTGATCCGAGAGTTATCCAGCCAGAGGGGACTTTGAC 797
QY 601 ATTGACATTTGGCCGTGTGATGACACAGTTGGGACCATGATGATGCTGCTATGAT 660
DB 798 ATTGACATTTGGCCGTGTGATGACACAGTTGGGACCATGATGATGCTGCTATGAT 857
QY 661 GATCAGAACTCGAGATTTGCTCATTTGTGGGCACTGGCAGCAACGCCCTGTACATGAG 720
DB 858 GATCAGAACTCGAGATTTGCTCATTTGTGGGCACTGGCAGCAACGCCCTGTACATGAG 917
QY 721 GAAATGCTCATTTGATGATGTGTGAGAGATGAGGGGCGCATGTGATCAACATGAG 780
DB 918 GAAATGCTCATTTGATGATGTGTGAGAGATGAGGGGCGCATGTGATCAACATGAG 977
QY 781 TGGGAGGCTTTGGGGAGAGGAGTACACTCAATGACATCCGAGATTTGACCGAGAG 840
DB 978 TGGGAGGCTTTGGGGAGAGGAGTACACTCAATGACATCCGAGATTTGACCGAGAG 1037
QY 841 ATGCATATGAGCTCGCTGACACCTTGGGAGAGAGCTGTTTGAAGATGATTAGCGGGATG 900
DB 1038 ATGCATATGAGCTCGCTGACACCTTGGGAGAGAGCTGTTTGAAGATGATTAGCGGGATG 1097
QY 901 TACATGGGGAGGCTGCTGAGGCTCATCTGTTGAAGATGGGCAAGGCGAGACTGTGTC 960
DB 1098 TACATGGGGAGGCTGCTGAGGCTCATCTGTTGAAGATGGGCAAGGCGAGACTGTGTC 1157
QY 961 CAAGGGAAGTACAGCCCAAGAACTCTTACCACTGCTCTTTCGAGACCAAGATGTCTCG 1020
DB 1158 CAAGGGAAGTACAGCCCAAGAACTCTTACCACTGCTCTTTCGAGACCAAGATGTCTCG 1217
QY 1021 GATATTGAAGAGATTAAGATGGAATGAGAAGGCTTACCAATCTGATGCGCTGGGT 1080
DB 1218 GATATTGAAGAGATTAAGATGGAATGAGAAGGCTTACCAATCTGATGCGCTGGGT 1277
QY 1081 CTGATTCATTTGAGAGAGATTTGTGTCAGCGACGCAAGTCTGCAGATTTGTGTCAGG 1140
DB 1278 CTGATTCATTTGAGAGAGATTTGTGTCAGCGACGCAAGTCTGCAGATTTGTGTCAGG 1337
QY 1141 CGCTTCGAGAGTCTGTGCGACCAACCTTGGCCGCGGTGCTGTGCGAATCAAGAGAGAC 1200
DB 1338 CGCTTCGAGAGTCTGTGCGACCAACCTTGGCCGCGGTGCTGTGCGAATCAAGAGAGAC 1397
QY 1201 AAGGCGAGAGAGCTTCCCTCCACCATGCGTGTGATGCTCGCTTACAAAGAAAT 1260
DB 1398 AAGGCGAGAGAGCTTCCCTCCACCATGCGTGTGATGCTCGCTTACAAAGAAAT 1457
QY 1261 CCCCATTGTCACACGCTCTCATTAAGGAGAGTGGAGGCTGCTCCGACATGATGTC 1320
DB 1458 CCCCATTGTCACACGCTCTCATTAAGGAGAGTGGAGGCTGCTCCGACATGATGTC 1517
QY 1321 CGCTTCTCTCGCTCTGAGAGATGACGAGGAGGGGCTGCTATGATGACGGGCTGCT 1380
DB 1518 CGCTTCTCTCGCTCTGAGAGATGACGAGGAGGGGCTGCTATGATGACGGGCTGCT 1577
QY 1381 TACGCTCTGCTGACCAACACCGGGGCGCCAGAAAGACCTTGAATCTGTAAGCTGAGC 1440
|||||

Db 1578 TACCTTGGCTGACCAACCGGGCCGCGAGAGACCTTGAGTCTCTAGAGCTGAGC 1637
 Qy 1441 CACGAGCAGCTTCTGGAGGTTAAGAGAAATGAGAGATGGAATATGAGCAGAGGCTGAGC 1500
 Db 1638 CACGAGCAGCTTCTGGAGGTTAAGAGAAATGAGAGATGGAATATGAGCAGAGGCTGAGC 1697
 Qy 1501 AAGAGAGCAGTGGGCTGCCCCCTGTGAAATGCTGCCCACTTACGTGTGCCACTCA 1560
 Db 1698 AAGAGAGCAGTGGGCTGCCCCCTGTGAAATGCTGCCCACTTACGTGTGCCACTCA 1757
 Qy 1561 GATGGCAGAGAAAGAGACTTCTGGCCCTTGGATCTGGAGAAACAACCTCCGGTTC 1620
 Db 1758 GATGGCAGAGAAAGAGACTTCTGGCCCTTGGATCTGGAGAAACAACCTCCGGTTC 1817
 Qy 1621 CTGCTGGCTGTGCTGTGCTTAATGAGCAACGAGCGGCTGAGATGATACAAAGATCTAC 1680
 Db 1818 CTGCTGGCTGTGCTGTGCTTAATGAGCAACGAGCGGCTGAGATGATACAAAGATCTAC 1877
 Qy 1681 TCCATCCCAAGAGAGTTATGATGCACTGGGAGAGAGCTCTTTCAGACACATTTGTCAG 1740
 Db 1878 TCCATCCCAAGAGAGTTATGATGCACTGGGAGAGAGCTCTTTCAGACACATTTGTCAG 1937
 Qy 1741 TGCATTCGCGAGCTTCTGGAGATACATGGGAGAGAGCGCTGCTGGCTTGGGTTTC 1800
 Db 1938 TGCATTCGCGAGCTTCTGGAGATACATGGGAGAGAGCGCTGCTGGCTTGGGTTTC 1997
 Qy 1801 ACATTTCTCTTCCCTTCCAGCAGACAGCCTAGACAGACATCTCTCAAGTGGACA 1860
 Db 1998 ACATTTCTCTTCCCTTCCAGCAGACAGCCTAGACAGACATCTCTCAAGTGGACA 2057
 Qy 1861 AAGGATTTCAAGGATCTGGCTGGAGGCTGAGAGATGTGGTCACTTGTCTGAAGAAAGCG 1920
 Db 2058 AAGGATTTCAAGGATCTGGCTGGAGGCTGAGAGATGTGGTCACTTGTCTGAAGAAAGCG 2117
 Qy 1921 ATTACCCGCGAGAGAGATTGACCTGATGTGGTCCGCTGGTGAATGACACAGTTGGG 1980
 Db 2118 ATTACCCGCGAGAGAGATTGACCTGATGTGGTCCGCTGGTGAATGACACAGTTGGG 2177
 Qy 1981 ACTATGATGACTTGTGGCTACGAAAGACCTTCACTGTGAAGTTGGCTCATTTGTTGGCACC 2040
 Db 2178 ACTATGATGACTTGTGGCTACGAAAGACCTTCACTGTGAAGTTGGCTCATTTGTTGGCACC 2237
 Qy 2041 GGAAGCAACCCCTGCTACATGGAAGAGATGCTAATGTGAGAGCTGGTGGAGGAGAGAG 2100
 Db 2238 GGAAGCAACCCCTGCTACATGGAAGAGATGCTAATGTGAGAGCTGGTGGAGGAGAGAG 2297
 Qy 2101 GGAAGCATGTGTGTCAACATGAGAGTGGGAGCATTTGGGGACAATGGCTCCCTGATGAC 2160
 Db 2298 GGAAGCATGTGTGTCAACATGAGAGTGGGAGCATTTGGGGACAATGGCTCCCTGATGAC 2357
 Qy 2161 TTGGGGACCGTGTGTTGATGTGTGCTGGAGATGAGACTTCTCTCAACCTGGCAAGAGAGG 2220
 Db 2358 TTGGGGACCGTGTGTTGATGTGTGCTGGAGATGAGACTTCTCTCAACCTGGCAAGAGAGG 2417
 Qy 2221 TTGGAGAAATGATCAAGCGGATGACTTGGAGAGATTGGCAACATTTCTCATGAT 2280
 Db 2418 TTGGAGAAATGATCAAGCGGATGACTTGGAGAGATTGGCAACATTTCTCATGAT 2477
 Qy 2281 TTACAGCAAGCGGGGCTGCTCTTCCAGAGCGCATCTCAGAGCGCTCAAGAGAGAGGA 2340
 Db 2478 TTACAGCAAGCGGGGCTGCTCTTCCAGAGCGCATCTCAGAGCGCTCAAGAGAGAGGA 2537
 Qy 2341 ATCTCTGAAACTAAGTTCCCTGCTCAAGATGAGAGAGAGAGCTTACGCTTACAGGTT 2400
 Db 2538 ATCTCTGAAACTAAGTTCCCTGCTCAAGATGAGAGAGAGAGCTTACGCTTACAGGTT 2597
 Qy 2401 CGTGCCATCTCTCGCCACCATAGGCGTGGAGAGCAGTGCATGACAGCATCATCTGAAG 2460
 Db 2598 CGTGCCATCTCTCGCCACCATAGGCGTGGAGAGCAGTGCATGACAGCATCATCTGAAG 2657
 Qy 2461 GAGGTGTGACTGTGTGGCCGCGGCTGCACAGCTCTGTGGCCAGGAGCATGGCCGCC 2520
 Db 2658 GAGGTGTGACTGTGTGGCCGCGGCTGCACAGCTCTGTGGCCAGGAGCATGGCCGCC 2717

Qy 2521 GTAGTGGACAGATTAAGAGAGAACCGTGGCTGGACACAACCCCAAGTACAGTGGCGTGC 2580
 Db 2718 GTAGTGGACAGATTAAGAGAGAACCGTGGCTGGACACAACCCCAAGTACAGTGGCGTGC 2777
 Qy 2581 GACGGGACTCTGTATTAAGCTTCATCTCAGCTTGTGCCAAGGCTAGCATGAGAGCGGTGAGA 2640
 Db 2778 GACGGGACTCTGTATTAAGCTTCATCTCAGCTTGTGCCAAGGCTAGCATGAGAGCGGTGAGA 2837
 Qy 2641 GATCTGGCTCCGAATATGATGACGTCTCTCTGGAATCCAGAGAGCGGAGTGGAGAGGA 2700
 Db 2838 GATCTGGCTCCGAATATGATGACGTCTCTCTGGAATCCAGAGAGCGGAGTGGAGAGGA 2897
 Qy 2701 GCAGCTCTCATCTGCGCGTGGCGCATCCGGAGGCTGGGAGAGATAG 2754
 Db 2898 GCAGCTCTCATCTGCGCGTGGCGCATCCGGAGGCTGGGAGAGATAG 2951

 RESULT 4
 AA00086
 ID AA00086 standard; cdna; 3635 bp.
 XX
 AC AA00086;
 XX
 DT 17-MAR-1998 (first entry)
 XX
 DE Rat hexokinase II encoding cDNA.
 XX
 KW Yeast; trehalose-6-phosphate synthase; hexokinase; ribozyme;
 KM stimulator; insulin; glucose; diabetes; ss.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT 198..2951
 FT CDS /tag= a
 FT /product= Hexokinase_II
 XX
 PD WO9726357-A1.
 XX
 PD 24-JUL-1997.
 XX
 PF 17-JAN-1997; 97WO-US00787.
 XX
 PR 19-JAN-1996; 96US-0588983.
 XX
 PA (BETA-) BETAGENE INC.
 PA (TEXAS) UNIV TEXAS SYSTEM.
 PI Han H, Newgard CB, Normington KD, Thigpen AE;
 XX
 XX WPI: 1997-385343/35.
 DR P-PSDB: AAW37429.
 XX
 PT Mammalian cell with reduced activity of low Km hexokinase - caused
 PT by specific ribozyme or stimulator of trehalose-6-phosphate
 PT production, particularly for in vitro or in vivo insulin production
 XX
 PS Claim 22: Page 186-189; 265pp; English.
 XX
 CC A mammalian cell has been developed comprising an effective amount of a
 CC low Km hexokinase-inhibitor selected from: (a) an agent that stimulates
 CC production of trehalose-6-phosphate (TPP); or (b) a ribozyme specific
 CC for low Km hexokinase. The present sequence encodes rat hexokinase II.
 CC The cells are particularly used to produce insulin, in response to
 CC glucose or other secretagogues, either in vitro or in vivo (for
 CC treating diabetes), but may also be used to produce many other
 CC therapeutic proteins, e.g., from a gene under control of the insulin
 CC promoter which is therefore responsive to glucose. Growth-inhibited
 CC cells (having altered hexokinase activity) are used to produce
 CC proteins, e.g., insulin or antibodies, in vivo or in vitro. Reduction in
 CC low Km hexokinase activity provides cells in which insulin secretion is
 CC induced at glucose concentrations closer to the normal range than in

1098 TACATGGGGAGCTGTCAGGCTCATCCCTGGTGAAGATGCCCAAGCAGAGCTGTTCTTC 1157

2178 ACTATGATGACTTGTGGCTACGAAGACCCTCACTGTGAAGTTGGCTCATTTCTTCCCAAGC 3337

```

OY 2041 GGAAGCAACCCCTGCTACATGAGAGATGCCATATGTGAGCTGTGAGCGAGAGAG 2100
    |||
DB 2238 GGAACCAACCCCTGCTACATGAGAGATGCCATATGTGAGCTGTGAGCGAGAGAG 2297
    |||
OY 2101 GAGCGATGTGTGTCAACATGAGTGGGAGCATTTGGGGCAATGGGCTGCTGATGAC 2160
    |||
DB 2298 GCACGATGTGTGTCAACATGAGTGGGAGCATTTGGGGCAATGGGCTGCTGATGAC 2357
    |||
OY 2161 TTGCGGACCGCTTTGATGTGTGTGATGATGATGATGATGATGATGATGATGATG 2220
    |||
DB 2358 TTGCGGACCGCTTTGATGTGTGTGATGATGATGATGATGATGATGATGATGATG 2417
    |||
OY 2221 TTGCGAGACATATCATGAGCGGATGATGATGATGATGATGATGATGATGATGATG 2280
    |||
DB 2418 TTGCGAGACATATCATGAGCGGATGATGATGATGATGATGATGATGATGATGATG 2477
    |||
OY 2281 TTGCGAGACCGGGGGGCTGCTTCCGAGGCGGATGATGATGATGATGATGATGATG 2340
    |||
DB 2478 TTGCGAGACCGGGGGGCTGCTTCCGAGGCGGATGATGATGATGATGATGATGATG 2537
    |||
OY 2341 ATCTCTGAACACTAAGTTCTCTCTCAGATAGAGAGCGGATGATGATGATGATGATG 2400
    |||
DB 2538 ATCTCTGAACACTAAGTTCTCTCTCAGATAGAGAGCGGATGATGATGATGATGATG 2597
    |||
OY 2401 CGTGCATCTCTGCGGACCTGAGAGAGCGGATGATGATGATGATGATGATGATGATG 2460
    |||
DB 2598 CGTGCATCTCTGCGGACCTGAGAGAGCGGATGATGATGATGATGATGATGATGATG 2657
    |||
OY 2461 GAGGTGTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520
    |||
DB 2658 GAGGTGTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2717
    |||
OY 2521 GTAGTGCACAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
    |||
DB 2718 GTAGTGCACAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2777
    |||
OY 2581 GAGGAGACTGTGTATAGCTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
    |||
DB 2778 GAGGAGACTGTGTATAGCTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2837
    |||
OY 2641 GATGTGCTCCGAATGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
    |||
DB 2838 GATGTGCTCCGAATGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2897
    |||
OY 2701 GCAGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2754
    |||
DB 2898 GCAGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2951
    |||

```

```

PA (UJO) UNIV JOHNS HOPKINS.
XX
XX Mathupala SP, Pedersen PL, Rempel A;
PI WPI: 1997-132643/12.
DR
XX
XX New transcription regulating fragments of hexokinase II DNA contg.
PT response element - and methods for diagnosis or treatment of
PT neoplasias that over-express hexokinase II and for regulating
PT glycolysis
PI
XX
XX Example 1: Fig 7, 104pp: English.
PS
XX
XX The present sequence represents the cDNA of the PCR-generated probe
CC used for the isolation of the promoter for Type II tumour hexokinase
CC from the AS-30D genomic library. AS-30D tumour Type II hexokinase
CC is a new isolated hexokinase II. DNA fragments from the promoter
CC region (shown in AAT78598), coding sequences and introns are able to
CC regulate transcription of a downstream open reading frame if they
CC contain at least one response element (transcription factor binding
CC site). A method has been produced for detecting a copy number greater than
CC 2 for the hexokinase II gene. The DNA fragments may be coupled to a
CC reporter gene and used to screen for potential drugs that affect
CC regulated transcription of tumour hexokinase II. Alternatively they may
CC be coupled to a toxic gene and used to treat cells that over-express
CC hexokinase II, such as those present in patients with cancer. They may
CC also be used in gene therapy to treat diabetes. The DNA fragments can
CC increase glycolysis in cells and express homologous or heterologous
CC protein. Probes of the DNA fragment are used in the method for
CC diagnosing a neoplasia that over-expresses hexokinase. The new response
CC elements are active only in tumours, not in normal cells.
XX
XX Sequence 299 BP; 59 A; 96 C; 76 G; 68 T; 0 other;
SQ
XX
XX Query Match 3.7%; Score 102; DB 18; Length 299;
XX Best local Similarity 100.0%; Pred. No. 2e-39;
XX Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGATGCGCTCGCATATGATGCGCTGCTTATTCACGAGCTCAACCAACCAAGTGAG 60
    |||
DB 198 ATGATGCGCTCGCATATGATGCGCTGCTTATTCACGAGCTCAACCAACCAAGTGAG 257
    |||
OY 61 AAGGTTGACCAATTTCTCTACCAATGCGTCTCTCAGATGAG 102
    |||
DB 258 AAGGTTGACCAATTTCTCTACCAATGCGTCTCTCAGATGAG 299
    |||

```

RESULT 5

AAT80522 ID AAT80522 standard; cDNA: 299 BP.

XX AAT80522:

XX 31-OCT-1997 (first entry)

XX AS-30D tumour Type II hexokinase probe.

XX Response element: 2-DNA: neoplasia; hexokinase II; glycolysis;

XX cancer; gene therapy; diabetes; tumour; rat; PCR;

XX polymerase chain reaction; ss.

XX Mus musculus.

XX W09704104-A2.

XX 06-FEB-1997.

XX 12-JUL-1996; 96WO-US11673.

XX 14-JUL-1995; 95US-0001199.

RESULT 6

AAT78598 ID AAT78598 standard; DNA: 5150 BP.

XX AAT78598:

XX 30-OCT-1997 (first entry)

XX AS-30D tumour Type II hexokinase 4.3kb proximal promoter region.

XX Response element: 2-DNA: neoplasia; hexokinase II; glycolysis;

XX cancer; gene therapy; diabetes; tumour; rat; ds.

XX Rattus rattus.

XX Key Location/Qualifiers

XX protein_bind

XX protein_bind

XX protein_bind

XX protein_bind

XX protein_bind

XX protein_bind

XX protein_bind

/tissue_type="Leukemia"
/note="Inverse PCR method
(http://genome2.ncicrf.gov/RTCGD)"

BASE COUNT 109 a 127 c 116 g 114 t
ORIGIN

Query Match 2.0%; Score 55; DB 17; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.2e-16;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATCGCCTCGCATATGATCGCTGCTTATTACGAGGCTCAACCAAAACCAAG 55
|||||
Db 55 ATGATCGCCTCGCATATGATCGCTGCTTATTACGAGGCTCAACCAAAACCAAG 1

RESULT 7

AW823261

LOCUS

DEFINITION AW823261 518 bp mRNA linear EST 17-MAY-2000
ur7la09.y1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:3155704 5,
similar to SW:HXK2_MOUSE O08528 HEXOKINASE TYPE II ;, mRNA
sequence.

ACCESSION

AW823261

VERSION

AW823261.1

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 518)

AUTHORS

NCI-CGAP

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MGI:1058460

Seq primer: -40RP from Gibco

High quality sequence stop: 418.

Location/Qualifiers

1..518

/organism="Mus musculus"

/strain="129,C57BL/6J,FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3155704"

/clone_lib="NCI_CGAP_Mam3"

/tissue_type="tumor, gross tissue"

/dev_stage="10 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."

BASE COUNT 125 a 133 c 142 g 118 t

ORIGIN

Query Match 1.9%; Score 51; DB 10; Length 518;

Best Local Similarity 100.0%; Pred. No. 1.4e-14;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 CATATGATCGCTGCTTATTACGAGGCTCAACCAAAACCAAGTCAGAG 63

|||||

Db 468 CATATGATCGCTGCTTATTACGAGGCTCAACCAAAACCAAGTCAGAG 518

```

RESULT 4
BQ951609
LOCUS
DEFINITION BQ951609 924 bp mRNA linear EST 21-AUG-2002
5', mRNA sequence.
ACCESSION BQ951609
VERSION BQ951609.1 GI:22367087
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13995 row: e column: 23
High quality sequence stop: 606.
Location/Qualifiers
1..924
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6468190"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 203 a 224 c 271 g 222 t 4 others
ORIGIN
FEATURES
source
Query Match 3.0%; Score 83; DB 14; Length 924;
Best Local Similarity 100.0%; Pred. No. 5.4e-31;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2620 GTCATGATGACAGCGTGAGAGATCGCTCGGAATGTGACGTGCTGCTGCAATCC 2679
|||||
Db 151 GTCATGATGACAGCGTGAGAGATCGCTCGGAATGTGACGTGCTGCTGCAATCC 210
|||||

QY 2680 GAGGACGCGAGTGGGAGGGAGC 2702
|||||
Db 211 GAGGACGCGAGTGGGAGGGAGC 233
|||||

RESULT 5
B1738353
LOCUS
DEFINITION B1738353 754 bp mRNA linear EST 20-SEP-2001
mRNA sequence.
ACCESSION B1738353
VERSION B1738353.1 GI:15715366
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

```

Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11937 row: m column: 03
High quality sequence start: 4
High quality sequence stop: 745.
Location/Qualifiers
1..754
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5368490"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES

source

```

BASE COUNT 181 a 190 c 219 g 164 t
ORIGIN
Query Match 2.6%; Score 71; DB 13; Length 754;
Best Local Similarity 100.0%; Pred. No. 7.8e-25;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATCGCTCGCATATGATCGCTGCTTATTCACGGAGCTCAACCAACCAAGTCCAG 60
|||||
Db 464 ATGATCGCTCGCATATGATCGCTGCTTATTCACGGAGCTCAACCAACCAAGTCCAG 523
|||||

QY 61 AAGGTTGACCA 71
|||||
Db 524 AAGGTTGACCA 534
|||||

```

```

RESULT 6
BH859067/c
LOCUS
DEFINITION BH859067 466 bp DNA linear GSS 08-JUL-2002
musculus genomic clone S5_319d, DNA sequence.
ACCESSION BH859067
VERSION BH859067.1 GI:21709888
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 466)
JOURNAL Suzukit, T., Shen, H., Akagi, K., Morse, H.C., Malley, J.D., Naiman, D.Q.,
COMMENT Retroviral tagging provides a potent cancer gene discovery tool in
the post-genome-sequence era
Nat. Genet., (2002) In press
Contact: Copeland NG
Mouse Cancer Genetics Program
National Cancer Institute
Bldg. 539, Rm. 229, Frederick, MD 21702-1201, USA
Tel: 301 846 1260
Fax: 301 846 6666
Email: copeland@ncifcrf.gov
Class: PCR with specific primers.
Location/Qualifiers
1..466
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="S5_319d"
/clone_lib="Mouse Retroviral Tagged Cancer Gene Database"
/sex="female"

```

```

Db 457 TTGCGACTTCCTGGAGTACATGGCATGAAGGCGGTGTCCTGCGCTTGGGTTTCACAT 398
      |||||||
QY 1805 TCCTCTCCCTTCCAGCAGACAGCTAGACCAGAGCATCTCCCTCAAGTGGACAAAGG 1864
      |||||||
Db 397 TCCTCTCCCTTCCAGCAGACAGCTAGACCAGAGCATCTCCCTCAAGTGGACAAAGG 338
      |||||||
QY 1865 GATTCAAGGCATCTGGCTGCGAGGGTGAGGATGTGGTCACTTGGCTGAAGGAAGCGGATTC 1924
      |||||||
Db 337 GATTCAAGGCATCTGGCTGCGAGGGTGAGGATGTGGTCACTTGGCTGAAGGAAGCGGATTC 278
      |||||||
QY 1925 ACCGGCAGAGAGAGTTTGACTCGATGTGGTTCGCGTGGTGAATGACACAGTTGGGACTA 1984
      |||||||
Db 277 ACCGGCAGAGAGAGTTTGACTCGATGTGGTTCGCGTGGTGAATGACACAGTTGGGACTA 218
      |||||||
QY 1985 TGATGACTTGTGGCTACGAAGACCTCAGTGTGAAGTTGGCCTCATTTGTTGGCACCGGAA 2044
      |||||||
Db 217 TGATGACTTGTGGCTACGAAGACCTCAGTGTGAAGTTGGCCTCATTTGTTGGCACCGGAA 158
      |||||||
QY 2045 GCAACGCTGCTACATGAAGAGATGGTAATGTGGAGCTGGTGGAGG 2092
      |||||||
Db 157 GCAACGCTGCTACATGAAGAGATGGTAATGTGGAGCTGGTGGAGG 110
      |||||||

RESULT 2
AA574507
LOCUS
DEFINITION
444 bp mRNA linear EST 02-SEP-1997
vm29p05.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA
clone IMAGE:991569 5', similar to gb:M75126 HEXOKINASE, TYPE I
(HUMAN); gb:J05277 Mouse hexokinase mRNA, complete cds (MOUSE);,
mRNA sequence.
ACCESSION
AA574507
VERSION
AA574507.1 GI:2349133
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 444)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicinep
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:563849
High quality sequence stop: 411.
FEATURES
Location/Qualifiers
1..444
/organism="Mus musculus"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:991569"
/clone_lib="Knowles Solter mouse blastocyst B1"
/tissue_type="blastocyst"
/dev_stage="embryo (pre-implantation)"
/lab_host="DH10B"
/notes="Organ: embryo; Vector: pSPORT; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally from mRNA prepared
from 800 blastocysts. Primer: SalI(dT):
5'-CGGTGCGACCTGCGACCGCTTTTCTTTT-3'. cDNAs were
cloned into the NotI/SalI sites of a pSPORT vector (Life
Technologies). Two different size selections: B1 (larger

```

```

inserts) and B3."
BASE COUNT 95 a 117 c 145 g 87 t
ORIGIN
Query Match 3.0%; Score 83; DB 9; Length 444;
Best Local Similarity 100.0%; Pred. No. 4.4e-31;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2620 GTCATGTCATGAGCGGTGAGAGATCTGGCTCCGAAATGTGACGTGCTCTCTGGAATCC 2679
      |||||||
Db 342 GTCATGTCATGAGCGGTGAGAGATCTGGCTCCGAAATGTGACGTGCTCTCTGGAATCC 401
      |||||||
QY 2680 GAGGACGGCAGTGGGAAGGAGC 2702
      |||||||
Db 402 GAGGACGGCAGTGGGAAGGAGC 424
      |||||||

RESULT 3
BG295479
LOCUS
DEFINITION
602392808F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4504920 5',
mRNA sequence.
ACCESSION
BG295479
VERSION
BG295479.1 GI:13057155
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 746)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabps-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL10377 row: 0 column: 01
High quality sequence stop: 736.
FEATURES
Location/Qualifiers
1..746
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4504920"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-gt primed.
Average insert size 3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 166 a 191 c 225 g 164 t
ORIGIN
Query Match 3.0%; Score 83; DB 12; Length 746;
Best Local Similarity 100.0%; Pred. No. 5.1e-31;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2620 GTCATGTCATGAGCGGTGAGAGATCTGGCTCCGAAATGTGACGTGCTCTCTGGAATCC 2679
      |||||||
Db 428 GTCATGTCATGAGCGGTGAGAGATCTGGCTCCGAAATGTGACGTGCTCTCTGGAATCC 487
      |||||||
QY 2680 GAGGACGGCAGTGGGAAGGAGC 2702
      |||||||
Db 488 GAGGACGGCAGTGGGAAGGAGC 510
      |||||||

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2003, 13:32:46.; Search time 3815 Seconds
(without alignments)
11691.316 Million cell updates/sec

Title: US-09-808-743a-1
Perfect score: 2754
Sequence: 1 atgatcgctcgcatatgat.....ggaggctggcgagagatag 2754

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 50

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rpd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	348	12.6	521	10 BE108528	BE108528 UI-R-CA0-
2	83	3.0	444	9 AA574507	AA574507 vm29b05.r
3	83	3.0	746	12 BG295479	BG295479 602392808
4	83	3.0	924	14 BQ951609	BQ951609 AGENCOURT
5	71	2.6	754	13 BI738353	BI738353 603361115
C 6	55	2.0	466	17 BH859067	BH859067 S5_319d_s

7 51 1.9 518 10 AW823261 ur71a09.y

ALIGNMENTS

RESULT 1
BE108528/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
.TITLE
JOURNAL
MEDLINE
COMMENT
1 (bases 1 to 521)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized hypothalamus library cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 1-32,
>POLY_A#Simple_repeat
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
Location/Qualifiers
1..521
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CA0-axf-c-10-0-UI"
/clone_lib="UI-R-CA0"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3b-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI. The UI-R-CA0
library is a subtracted library derived from the following
tissues: thalamus, cerebellum, hypothalamus, medulla, pons
, midbrain, cerebral cortex, corpus striatum, testis, and
hippocampus. For a detailed description of the library
from which this clone was derived, please visit our web
site at ratest.eng.uiowa.edu. The subtraction has been
previously described in (Bonaldo, Lennon and Soares,
Genome Research 6:791-806, 1996)
TAG_LIB=UI-R-CA0
TAG_TISSUE=hypothalamus
TAG_SEQ=GATGC"

BASE COUNT 114 a 157 c 112 g 138 t
ORIGIN
Query Match 12.6%; Score 348; DB 10; Length 521;
Best Local Similarity 100.0%; Pred. No. 1.3e-167;
Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1745 TTGCGGACTTCTGTGGATACATGCGGCATGAGGGCGTGTCCCTTGGGTTTCACAT 1804

Qy	901	TACATGGGGAGCTGGTCTAGCCTCATCTCGTGAAGATGCCAAGCAGAGCTGTGTTC	961
Db	901	TACATGGGGAGCTGGTCTAGCCTCATCTCGTGAAGATGCCAAGCAGAGCTGTGTTC	961
Qy	961	CAAGGGAACCTACGCCAGAACTCCTTTACCACCTGGCTCCCTTCGAGACCAGAACTGTCGC	1021
Db	961	CAAGGGAACCTACGCCAGAACTCCTTTACCACCTGGCTCCCTTCGAGACCAGAACTGTCGC	1021
Qy	1021	GATATTGAAGAGGATAGGATGGAAATCGAAGGCCTACCAATCTGATGCCTGGGT	1081
Db	1021	GATATTGAAGAGGATAGGATGGAAATCGAAGGCCTACCAATCTGATGCCTGGGT	1081
Qy	1081	CTGAATCCATTGCAAGGAGGATTGTGTGGCCACCACCCGAATCTGCCAGATTCTGTCCACG	1141
Db	1081	CTGAATCCATTGCAAGGAGGATTGTGTGGCCACCACCCGAATCTGCCAGATTCTGTCCACG	1141
Qy	1141	CGCTCGGCAGTCTGTGCGAGCCACCCTGGCCGCGGTCTGTGGCGAATCAAGAGAAC	1201
Db	1141	CGCTCGGCAGTCTGTGCGAGCCACCCTGGCCGCGGTCTGTGGCGAATCAAGAGAAC	1201
Qy	1201	AAGGCGAGGAGGCACCTTCCTCCACCATCGGTGTCGATGGCTCGTCTACAAGAAAACAT	1261
Db	1201	AAGGCGAGGAGGCACCTTCCTCCACCATCGGTGTCGATGGCTCGTCTACAAGAAAACAT	1261
Qy	1261	CCCAATTTGCCAAGCGTCTCCATAAGGCAGTGAAGAGCCTGGTCCCGACTGTGATGTC	1321
Db	1261	CCCAATTTGCCAAGCGTCTCCATAAGGCAGTGAAGAGCCTGGTCCCGACTGTGATGTC	1321
Qy	1321	CGCTTCCTCCGCTCTGAGGATGCGACGCGAAGGGGCTGTATGGTGACGGCGTGCGT	1381
Db	1321	CGCTTCCTCCGCTCTGAGGATGCGACGCGAAGGGGCTGTATGGTGACGGCGTGCGT	1381
Qy	1381	TACCGTCTGGCTGACCAACACCGGGGCCGCGAAGACCCCTGGAGTCTCTGAAGCTGAGC	1441
Db	1381	TACCGTCTGGCTGACCAACACCGGGGCCGCGAAGACCCCTGGAGTCTCTGAAGCTGAGC	1441
Qy	1441	CACGACGAGCTCTGAGGTTTAAAGAAATGAAGTGGGAAATGGAGCAGGGTCTGAGC	1501
Db	1441	CACGACGAGCTCTGAGGTTTAAAGAAATGAAGTGGGAAATGGAGCAGGGTCTGAGC	1501
Qy	1501	AAGGAGCGCATCGGTCGCGCTCTGGAAGATGCTGCCACCTACGTCGTGTGCCACCTCCA	1561
Db	1501	AAGGAGCGCATCGGTCGCGCTCTGGAAGATGCTGCCACCTACGTCGTGTGCCACCTCCA	1561
Qy	1561	GATGGCACAGAAAGAGACATCTTGGCCCTTGGATCTTGGAGGAACAACTCCCGGGTC	1621
Db	1561	GATGGCACAGAAAGAGACATCTTGGCCCTTGGATCTTGGAGGAACAACTCCCGGGTC	1621
Qy	1621	CTGCTGTGCGTGTGGGTAAATGCCAAGCGAGGGCGCTGGAGATCATTAACAAGATCTAC	1681
Db	1621	CTGCTGTGCGTGTGGGTAAATGCCAAGCGAGGGCGCTGGAGATCATTAACAAGATCTAC	1681
Qy	1681	TCCATCCACAGAGGTTATGCATGGCACTGGGGAAGAGCTCTTCGACCCACATTTGCCAG	1741
Db	1681	TCCATCCACAGAGGTTATGCATGGCACTGGGGAAGAGCTCTTCGACCCACATTTGCCAG	1741
Qy	1741	TGCATTGCGGACTTCTTGAGTACATGGGCATGAAGGCGGTGCCCTGGCTTGGGTTTC	1801
Db	1741	TGCATTGCGGACTTCTTGAGTACATGGGCATGAAGGCGGTGCCCTGGCTTGGGTTTC	1801
Qy	1801	ACATCTCCTCTCCCTTGGCAGCAGAACCCCTAGACAGACATCCCTCAAGTGGACA	1861
Db	1801	ACATCTCCTCTCCCTTGGCAGCAGAACCCCTAGACAGACATCCCTCAAGTGGACA	1861
Qy	1861	AAGGGATTCAAGGCATCTGGCTGCGAGGTTGAGGATGTGGTCACTTGTGAAGGAAGCG	1921
Db	1861	AAGGGATTCAAGGCATCTGGCTGCGAGGTTGAGGATGTGGTCACTTGTGAAGGAAGCG	1921
Qy	1921	ATTACCGGCGAGAGGAGTTTCACTTGGATGTGGTTCGCGTGGTGAATGACACAGTTGGG	1981
Db	1921	ATTACCGGCGAGAGGAGTTTCACTTGGATGTGGTTCGCGTGGTGAATGACACAGTTGGG	1981
Qy	1981	ACTATGATGACTTGTGGCTTACGAAGACCCCTCACTGTGAAGTTGGCCTCATTTGTGGCACC	2041

[illegible]

Search completed: May 27, 2003, 19:22:56
Job time : 362 secs

RESULT 1
US-09-808-743-1
; Sequence 1, Application US/09808743
; Patent No. US20020068711a1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: PEDERSEN, Peter
; APPLICANT: MATHUPALA, Saroj
; TITLE OF INVENTION: ARREST OF PROLIFERATION OF HIGHLY GLYCOLYTIC TUMORS
; FILE REFERENCE: JH01720-1
; CURRENT APPLICATION NUMBER: US/09/808,743
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/189,222

Search completed: May 27, 2003, 17:13:56
Job time : 107 secs

```
Db 1818 CTGCTGGTGGCTGTGCGTAATGTGCAAGCGAGGGCGTGGAGATGCATAAACAAGATCTAC 1877
Qy 1681 TCCATCCACAGAGAGGTTATGCATGGCACTGGGGAAGAGCTCTTGGAAACACATTTGCCAG 1740
Db 1878 TCCATCCACAGAGAGGTTATGCATGGCACTGGGGAAGAGCTCTTGGAAACACATTTGCCAG 1937
Qy 1741 TGCATTGCGGACTTCCCTGGAGTACATGGGCATGAAGGGGCTGCTCCCTGGCTTTGGGTTTC 1800
Db 1938 TGCATTGCGGACTTCCCTGGAGTACATGGGCATGAAGGGGCTGCTCCCTGGCTTTGGGTTTC 1997
Qy 1801 ACATTCTCTCTTCCCTGGCCAGAGAACAGCCTAGACCAGAGATCCTCTCAAGTGGACA 1860
Db 1998 ACATTCTCTCTTCCCTGGCCAGAGAACAGCCTAGACCAGAGATCCTCTCAAGTGGACA 2057
Qy 1861 AAGGGATTCAAGGCATCTCGCTGCGAGGTGAGGATGTGGTCACCTTGGTGAAGTGGAG 1920
Db 2058 AAGGGATTCAAGGCATCTGGCTGCGAGGTGAGGATGTGGTCACCTTGGTGAAGTGGAG 2117
Qy 1921 ATTCACCGCGAGAGAGTTTGACCTGGATGTGGTGGCTGGTGAATGACACAGTTGGG 1980
Db 2118 ATTCACCGCGAGAGAGTTTGACCTGGATGTGGTGGCTGGTGAATGACACAGTTGGG 2177
Qy 1981 ACTATGATGACTTGTGGCTACGAAGACCTCACTGTGAAGTTGGCTCATTTGTTGGCACC 2040
Db 2178 ACTATGATGACTTGTGGCTACGAAGACCTCACTGTGAAGTTGGCTCATTTGTTGGCACC 2237
Qy 2041 GGAACAACGCCGTGTACATGGAAGAGATGCGTAATGTGAGCTGGTGGACGGAGAGGAG 2100
Db 2238 GGAACAACGCCGTGTACATGGAAGAGATGCGTAATGTGAGCTGGTGGACGGAGAGGAG 2297
Qy 2101 GGACGATGTGTCAACATGAGTGGGAGCATTTGGGGACAAATGGCTGGCTGGATGAC 2160
Db 2298 GGACGATGTGTCAACATGAGTGGGAGCATTTGGGGACAAATGGCTGGCTGGATGAC 2357
Qy 2161 TTGCGGACCGGTTTGATGTGCTGTGGATGAGCTTTCTCAACCTGGCAACAGAGG 2220
Db 2358 TTGCGGACCGGTTTGATGTGCTGTGGATGAGCTTTCTCAACCTGGCAACAGAGG 2417
Qy 2221 TTCGAGAAGATGATCAGCGCATGTACTTGGGAGAGATTGTGGCAACATTTCTCATCGAT 2280
Db 2418 TTCGAGAAGATGATCAGCGCATGTACTTGGGAGAGATTGTGGCAACATTTCTCATCGAT 2477
Qy 2281 TTCAGAAAGCGGGGCTGCTCTCCGAGGCGCATCTCAGAGCGCCCTCAAGACAAGGGA 2340
Db 2478 TTCAGAAAGCGGGGCTGCTCTCCGAGGCGCATCTCAGAGCGCCCTCAAGACAAGGGA 2537
Qy 2341 ATCTCTGAACCTAAGTTCCCTGTCTCAGATAGAGGACATGCGCTAGCCCTGTACAGGTT 2400
Db 2538 ATCTTGAACCTAAGTTCCCTGTCTCAGATAGAGGACATGCGCTAGCCCTGTACAGGTT 2597
Qy 2401 CGTGCCATCTCGGCCACCTAGGCTGGAGACACGTGGCATGACAGCATCATCTGTGAAG 2460
Db 2598 CGTGCCATCTCGGCCACCTAGGCTGGAGACACGTGGCATGACAGCATCATCTGTGAAG 2657
Qy 2461 GAGGTGTCACTGTGTTGCCCGGGCGGTGCACAGCTGTGGCGCAGGCATGGCCGCC 2520
Db 2658 GAGGTGTCACTGTGTTGCCCGGGCGGTGCACAGCTGTGGCGCAGGCATGGCCGCC 2717
Qy 2521 GTAGTGACAAGATAAGAGAAACCGTGGGCTGGCAACCCCAAGTGCAGTGGCGCTG 2580
Db 2718 GTAGTGACAAGATAAGAGAAACCGTGGGCTGGCAACCCCAAGTGCAGTGGCGCTG 2777
Qy 2581 GACGGACTCTGTATAGCTTCACTCTCACTTTGCCAAGTGCATGATGAGACGGTGAGA 2640
Db 2778 GACGGACTCTGTATAGCTTCACTCTCACTTTGCCAAGTGCATGATGAGACGGTGAGA 2837
Qy 2641 GATCTGGCTCCGAAATGTGACGTGCTCTTCCCTGGAAATCCGAGAGCGGAGTGGGAAGGA 2700
Db 2838 GATCTGGCTCCGAAATGTGACGTGCTCTTCCCTGGAAATCCGAGAGCGGAGTGGGAAGGA 2897
Qy 2701 GCAGCTCTCATCACTGCCCTGGGCTGCCGCATCCGGGAGGCTGGGCAGAGATAG 2754
Db 2898 GCAGCTCTCATCACTGCCCTGGGCTGCCGCATCCGGGAGGCTGGGCAGAGATAG 2951
```

CORRESPONDENCE ADDRESS:		
ADDRESSEE:	Arnold, White & Durkee	
STREET:	P.O. Box 4433	
CITY:	Houston	
STATE:	TX	
COUNTRY:	US	
ZIP:	77210	
COMPUTER READABLE FORM:		
MEDIUM TYPE:	Floppy disk	
COMPUTER:	IBM PC compatible	
OPERATING SYSTEM:	PC-DOS/MS-DOS	
SOFTWARE:	PatentIn Release #1.0, Version #1.30	
CURRENT APPLICATION DATA:		
APPLICATION NUMBER:	US/08/588,976	
FILING DATE:	Concurrently herewith	
CLASSIFICATION:		435
ATTORNEY/AGENT INFORMATION:		
NAME:	Fussey, Shelley P.M.	
REGISTRATION NUMBER:	39,458	
REFERENCE/DOCKET NUMBER:	UTSD:481/FUS	
TELECOMMUNICATION INFORMATION:		
TELEPHONE:	(512) 418-3000	
TELEFAX:	(512) 474-7577	
TELEX:	n/a	
INFORMATION FOR SEQ ID NO:		15:
SEQUENCE CHARACTERISTICS:		
LENGTH:	3635 base pairs	
TYPE:	nucleic acid	
STRANDEDNESS:	single	
TOPOLOGY:	linear	
US-08-588-976-15		
Query Match		92.6%; Score 2550; DB 2; Length 3635;
Best Local Similarity		99.9%; Pred. No. 0;
Matches 2750; Conservative		0; Mismatches 4; Indels 0; Gaps 0;
QY	1	ATGATCGCTCGCATATGATCGCTGCTTATTACAGGAGCTCAACAAACCAAGTGCGAG 60
Db	198	ATGATCGCTCGCATATGATCGCTGCTTATTACAGGAGCTCAACAAACCAAGTGCGAG 257
QY	61	AAGTTGACCAATTTCTACACACATGCGTCTCAGATGAGACCTTCTGGAGATTCT 120
Db	258	AAGTTGACCAATTTCTACACACATGCGTCTCAGATGAGACCTTCTGGAGATTCT 317
QY	121	AGCGGTTCCGGAAGGAGATGGAGAAGGCTAGGAGCTACACGACACCTACAGCAGCT 180
Db	318	AGCGGTTCCGGAAGGAGATGGAGAAGGCTAGGAGCTACACGACACCTACAGCAGCT 377
QY	181	GTGAAATGTTGCTACCTTTGTGAGGTCAACTCCGGATGGGACAGAAATGGGGAGTTC 240
Db	378	GTGAAATGTTGCTACCTTTGTGAGGTCAACTCCGGATGGGACAGAAATGGGGAGTTC 437
QY	241	CTGCGCTCTGGATCTTGAGGAACCAACTTCCGTGCTCCGAGTAAGGGTGACGGACAAT 300
Db	438	CTGCGCTCTGGATCTTGAGGAACCAACTTCCGTGCTCCGAGTAAGGGTGACGGACAAT 497
QY	301	GGCTCCAGAGTGAGATGGAGAACCAGATCTACGCCATCTTGAGGACATCATCGCG 360
Db	498	GGCTCCAGAGTGAGATGGAGAACCAGATCTACGCCATCTTGAGGACATCATCGCG 557
QY	361	GGCAGTGGAAACCCAGCTGTTTGACACATCGCGAAATGCCCTGGCCAACTTCATGGACAAG 420
Db	558	GGCAGTGGAAACCCAGCTGTTTGACACATCGCGAAATGCCCTGGCCAACTTCATGGACAAG 617
QY	421	CTACAAATCAAGAGAGAAGCTCCCTCTGGGTTTACCTTCTGTTTCCCTCGCCACAG 480
Db	618	CTACAAATCAAGAGAGAAGCTCCCTCTGGGTTTACCTTCTGTTTCCCTCGCCACAG 677
QY	481	ACAAACTGGATGAGAGTTTTTGGTCTCTGCTGGACTAAGGGGTTCAAGTCCAGTGGCGTG 540
Db	678	ACAAACTGGATGAGAGTTTTTGGTCTCTGCTGGACTAAGGGGTTCAAGTCCAGTGGCGTG 737
QY	541	GAAGSCAGAGATGTTGGTGGACCTGATCCGGAAGGTTATCCAGCCAGAGGGGACTTTGAC 600
Db	738	GAAGSCAGAGATGTTGGTGGACCTGATCCGGAAGGTTATCCAGCCAGAGGGGACTTTGAC 797
QY	601	ATTGACATTTGGCCCTGGTGAATGACACAGTTGGGACCATGATGACTTGTGCTGCTATGAT 660
Db	798	ATTGACATTTGGCCCTGGTGAATGACACAGTTGGGACCATGATGACTTGTGCTGCTATGAT 857
QY	661	GATCAGAAGCTGCGAGATTTGGTCTCATTGTGGCACTGGCAGCAACCCCTGCTACATGGAG 720
Db	858	GATCAGAAGCTGCGAGATTTGGTCTCATTGTGGCACTGGCAGCAACCCCTGCTACATGGAG 917
QY	721	GAAATGCGTCATATTGACATGTTGGAGGAGATGAGGGGGCCATGTGCATCAACATGGAG 780
Db	918	GAAATGCGTCATATTGACATGTTGGAGGAGATGAGGGGGCCATGTGCATCAACATGGAG 977
QY	781	TGGGGAGCCTTTGGGACGACGGTACACTCAATGACATCCGAACCGAGTTTGGACCGAGAG 840
Db	978	TGGGGAGCCTTTGGGACGACGGTACACTCAATGACATCCGAACCGAGTTTGGACCGAGAG 1037
QY	841	ATCGACATGGGCTCGCTGAACCCCTGGGAAGCAGCTGTTTGAAGATGATTAAGCGGATG 900
Db	1038	ATCGACATGGGCTCGCTGAACCCCTGGGAAGCAGCTGTTTGAAGATGATTAAGCGGATG 1097
QY	901	TACATGGGGAGCTGTTGAGGCTCATCTGTTGAAGATGCGCAAGGCAGAGCTGTTGTTTC 960
Db	1098	TACATGGGGAGCTGTTGAGGCTCATCTGTTGAAGATGCGCAAGGCAGAGCTGTTGTTTC 1157
QY	961	CAAGGAAACTCAGCCCGACAGACTCCTTACCACCTGGCTCTTCGAGACCAAGATGTCG 1020
Db	1158	CAAGGAAACTCAGCCCGACAGACTCCTTACCACCTGGCTCTTCGAGACCAAGATGTCG 1217
QY	1021	GATATTGAAGAGATAAGGATGGAATCGAAGAGCCTTACCAATCTTGATGCGCTCGGT 1080
Db	1218	GATATTGAAGAGATAAGGATGGAATCGAAGAGCCTTACCAATCTTGATGCGCTCGGT 1277
QY	1081	CTGAATCCATTGCAAGGAGATTGTTGGCCACGACCAAGATCTGCCAGATTGTGTCACG 1140
Db	1278	CTGAATCCATTGCAAGGAGATTGTTGGCCACGACCAAGATCTGCCAGATTGTGTCACG 1337
QY	1141	CGCTCGGCGAGTCTGTGCGCAGCCACCTGCGCGGGTGTGTGGCGAATCAAGAGAGAC 1200
Db	1338	CGCTCGGCGAGTCTGTGCGCAGCCACCTGCGCGGGTGTGTGGCGAATCAAGAGAGAC 1397
QY	1201	AAGGCGAGGAGGAGCTTCCGCTCCACCATCGGTGCGATGCTGCTTACCAAGAAACAT 1260
Db	1398	AAGGCGAGGAGGAGCTTCCGCTCCACCATCGGTGCGATGCTGCTTACCAAGAAACAT 1457
QY	1261	CCCCATTTGGCAACGCTCTCCATTAAGCAGTGAAGAGGCTGGTGGCCGAGCTGTGATGTC 1320
Db	1458	CCCCATTTGGCAACGCTCTCCATTAAGCAGTGAAGAGGCTGGTGGCCGAGCTGTGATGTC 1517
QY	1321	CGCTTCTCCGCTCTGAGGATGGCAGGGGCTGCTATGCTGACGCGGCTGGCT 1380
Db	1518	CGCTTCTCCGCTCTGAGGATGGCAGGGGCTGCTATGCTGACGCGGCTGGCT 1577
QY	1381	TACCGTCTGGCTGACCAACACCGCGCCCAAGAACCTTGGAGTCTCTGAAGCTGAGC 1440
Db	1578	TACCGTCTGGCTGACCAACACCGCGCCCAAGAACCTTGGAGTCTCTGAAGCTGAGC 1637
QY	1441	CACGAGCAGCTTCTGGAGGTTAAGAGAGAATGAAGTGGAAATGAAGCAGGCTGTGAGC 1500
Db	1638	CACGAGCAGCTTCTGGAGGTTAAGAGAGAATGAAGTGGAAATGAAGCAGGCTGTGAGC 1697
QY	1501	AAGGAGACGATGCGGCTCGCCCTGTGAAGATGCTGCCACCTTACGTGTGTCGCTTCCA 1560
Db	1698	AAGGAGACGATGCGGCTCGCCCTGTGAAGATGCTGCCACCTTACGTGTGTCGCTTCCA 1757
QY	1561	GATGGCAGAGAGAAAGAGACTTCTTGGCTTGGATCTTTGGAGGAACAACCTTCCGGGTC 1620
Db	1758	GATGGCAGAGAGAAAGAGACTTCTTGGCTTGGATCTTTGGAGGAACAACCTTCCGGGTC 1817
QY	1621	CTGCTGGTGGCTGCTGCTAATGGCAAGCGGAGGGGCTGGAGATGCATCAAGATCTAC 1680

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P. M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ. ID NO. 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 3635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-15

Query Match 92.6% Score 2550; DB 2; Length 3635;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2750; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGATCGCTCGCATATGATCGCTTCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG 60
Db 198 ATGATCGCTCGCATATGATCGCTTCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG 257
Qy 61 AAGTTGACCAATTTCTTACCACATGCTCTCAGATGAGCCCTTCTGGAGATTCT 120
Db 258 AAGTTGACCAATTTCTTACCACATGCTCTCAGATGAGCCCTTCTGGAGATTCT 317
Qy 121 AGCGGTTCCGGAAGGAGATGGAGAAAGGCTTAGGAGTACACACACCTACAGCAGCT 180
Db 318 AGCGGTTCCGGAAGGAGATGGAGAAAGGCTTAGGAGTACACACACCTACAGCAGCT 377
Qy 181 GTGAAATTTCCCTACCTTTGTGAGTCAACTCCGATGGGAGAGACATGGGAGTTC 240
Db 378 GTGAAATTTCCCTACCTTTGTGAGTCAACTCCGATGGGAGAGACATGGGAGTTC 437
Qy 241 CTGGCTCTCGATCTTGGAGAACCAACTTCCGTCGAGTAAGGAGTACGGAGCAAT 300
Db 438 CTGGCTCTCGATCTTGGAGAACCAACTTCCGTCGAGTAAGGAGTACGGAGCAAT 497
Qy 301 GGCCTCCAGAGATGGAGATGGAGAACCAACTTCCGTCGAGTAAGGAGTACGGAGCAAT 360
Db 498 GGCCTCCAGAGATGGAGATGGAGAACCAACTTCCGTCGAGTAAGGAGTACGGAGCAAT 557
Qy 361 GGCAGTGGAAACCCAGCTGTTGACCAATCGCCGATGCTGGCCAACTTCATGGACAAG 420
Db 558 GGCAGTGGAAACCCAGCTGTTGACCAATCGCCGATGCTGGCCAACTTCATGGACAAG 617
Qy 421 CTACAAATCAAGAGAGAGCTCCCTCTGGGTTTCACTTCTCGTTCCCTGCCACAG 480
Db 618 CTACAAATCAAGAGAGAGCTCCCTCTGGGTTTCACTTCTCGTTCCCTGCCACAG 677
Qy 481 ACAAACTGGATCAGAGTCTTGGTCTCGTGAGTAAAGGGTTCAAGTCCAGTGGCGTG 540
Db 678 ACAAACTGGATCAGAGTCTTGGTCTCGTGAGTAAAGGGTTCAAGTCCAGTGGCGTG 737
Qy 541 GAAGGAGAGATGTGGTGGACCTGATCCGGAAGTTTATCCAGCGAGAGGGGACTTTGAC 600
Db 738 GAAGGAGAGATGTGGTGGACCTGATCCGGAAGTTTATCCAGCGAGAGGGGACTTTGAC 797
Qy 601 ATTGACATTTGGCCGTGGTGAATGACACAGTTGGGACCATGATGACTTCTGGCTATGAT 660
Db 798 ATTGACATTTGGCCGTGGTGAATGACACAGTTGGGACCATGATGACTTCTGGCTATGAT 857
Qy 661 GATCAGAACTCGGAGATTGGTCTCATTTGTGGCACTGGCAGCAACGCCCTACATGGAG 720
Db 858 GATCAGAACTCGGAGATTGGTCTCATTTGTGGCACTGGCAGCAACGCCCTACATGGAG 917

OM nucleic - nucleic search, using sw model
Run on: May 27, 2003, 13:52:56 ; Search time 101 Seconds
(without alignments)
8562.255 Million cell updates/sec

Title: US-09-808-743A-1
Perfect score: 2754
Sequence: 1 atgacgctcgcatatg.....gggaggctggcagagatag 2754

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0
Searched: 441362 seqs, 15333881 residues

Word size : 50
Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Listing first 1000 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Match Length	Description
1	2550	92.6 3635	2 US-08-588-983-15 Sequence 15, Appl
2	2550	92.6 3635	2 US-08-588-976-15 Sequence 15, Appl

ALIGNMENTS
RESULT 1
US-08-588-983-15
Sequence 15, Application US/08588983
Patent No. 5854067
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
for Inhibiting Hexokinase
TITLE OF INVENTION: for Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

FT      protein_bind      /note= "Transcription factor binding site"
FT      3162..3167
FT      /tag= an
FT      /label= Myb
FT      /note= "Transcription factor binding site"
FT      3223..3228
FT      /tag= ao
FT      /label= SRE
FT      /note= "Transcription factor binding site"
FT      3256..3267
FT      /tag= ap
FT      /label= Sp-1, Ap-2, SRE
FT      /note= "Transcription factor binding site"
FT      3306..3313
FT      /tag= aq
FT      /label= c/ebp
FT      /note= "Transcription factor binding site"
FT      3352..3360
FT      /tag= ar
FT      /label= NF-IL6
FT      /note= "Transcription factor binding site"
FT      3466..3483
FT      /tag= as
FT      /rpt_type= DIRECT
FT      /label= II
FT      3492..3509
FT      /tag= at
FT      /rpt_type= DIRECT
FT      /label= III
FT      3507..3513
FT      /tag= au

```

```

Query Match      2.4%; Score 65; DB 18; Length 5150;
Best Local Similarity 100.0%; Pred. No. 2.7e-21;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1  ATGATCGCCTCGCATATGATGCCCTGCTATTTCACGGAGCTCAACCAAAACCAAGTGCAG 60
Db      4831 ATGATCGCCTCGCATATGATGCCCTGCTATTTCACGGAGCTCAACCAAAACCAAGTGCAG 60
OY      61  AAGGT 65
Db      4891 AAGGT 4895

```

```

Search completed: May 27, 2003, 14:07:54
Job time : 624 secs

```

Wed May 28 08:37:51 2003

```
FT FT      /*tag= c
FT FT      /label= p53
FT FT      /note= "Transcription factor binding site"
FT FT      208..220
FT FT      /*tag= d
FT FT      /label= HNF-5,AP-3,c/ebp
FT FT      /note= "Transcription factor binding site"
FT FT      231..240
FT FT      /*tag= e
FT FT      /label= IRE
FT FT      /note= "Transcription factor binding site"
FT FT      322..327
FT FT      /*tag= f
FT FT      /label= Myb
FT FT      /note= "Transcription factor binding site"
FT FT      453..459
FT FT      /*tag= g
FT FT      /label= HNF-5
FT FT      /note= "Transcription factor binding site"
FT FT      504..523
FT FT      /*tag= h
FT FT      /label= Pea-3,SRE,Ap-2
FT FT      /note= "Transcription factor binding site"
FT FT      529..559
FT FT      /*tag= i
FT FT      /rpt_type= DIRECT
FT FT      /note= "Motif with the potential to form Z-DNA
FT FT      structure"
FT FT      605..610
FT FT      /*tag= j
FT FT      /label= GLRE,MYC
FT FT      /note= "Transcription factor binding site"
FT FT      640..647
FT FT      /*tag= k
FT FT      /label= c/ebp
FT FT      /note= "Transcription factor binding site"
FT FT      723..728
FT FT      /*tag= l
FT FT      /label= Pea-3
FT FT      /note= "Transcription factor binding site"
FT FT      741..746
FT FT      /*tag= m
FT FT      /label= Pea-3
FT FT      /note= "Transcription factor binding site"
FT FT      899..906
FT FT      /*tag= n
FT FT      /label= Ap-1
FT FT      /note= "Transcription factor binding site"
FT FT      1037..1043
FT FT      /*tag= o
FT FT      /label= HNF-5
FT FT      /note= "Transcription factor binding site"
FT FT      1076..1084
FT FT      /*tag= p
FT FT      /label= Sp-1
FT FT      /note= "Transcription factor binding site"
FT FT      1112..1117
FT FT      /*tag= q
FT FT      /label= Pea-3
FT FT      /note= "Transcription factor binding site"
FT FT      1414..1419
FT FT      /*tag= r
FT FT      /label= pPAR
FT FT      /note= "Transcription factor binding site"
FT FT      1634..1640
FT FT      /*tag= s
FT FT      /label= Ap-1
FT FT      /note= "Transcription factor binding site"
FT FT      1816..1824
FT FT      /*tag= t
FT FT      /label= c/ebp,NF-IL6
FT FT      /note= "Transcription factor binding site"
FT FT      2045..2051
FT FT      /tag= u
FT FT      /label= Ap-1
FT FT      /note= "Transcription factor binding site"
FT FT      2090..2099
FT FT      /*tag= v
FT FT      /rpt_type= DIRECT
FT FT      /label= I
FT FT      2133..2142
FT FT      /*tag= w
FT FT      /rpt_type= DIRECT
FT FT      /label= I
FT FT      2144..2152
FT FT      /*tag= x
FT FT      /label= Sp-1
FT FT      /note= "Transcription factor binding site"
FT FT      2165..2171
FT FT      /*tag= y
FT FT      /label= HNF-5
FT FT      /note= "Transcription factor binding site"
FT FT      2325..2332
FT FT      /*tag= z
FT FT      /label= Ap-2
FT FT      /note= "Transcription factor binding site"
FT FT      2402..2409
FT FT      /*tag= aa
FT FT      /label= Ap-2
FT FT      /note= "Transcription factor binding site"
FT FT      2413..2420
FT FT      /*tag= ab
FT FT      /label= Ap-1
FT FT      /note= "Transcription factor binding site"
FT FT      2593..2636
FT FT      /*tag= ac
FT FT      /rpt_type= DIRECT
FT FT      2757..2766
FT FT      /*tag= ad
FT FT      /label= p53
FT FT      /note= "Transcription factor binding site"
FT FT      2776..2784
FT FT      /*tag= ae
FT FT      /label= NF-IL6,Ap-1
FT FT      /note= "Transcription factor binding site"
FT FT      2866..2873
FT FT      /*tag= af
FT FT      /label= Ap-2
FT FT      /note= "Transcription factor binding site"
FT FT      2908..2919
FT FT      /*tag= ag
FT FT      /label= Myb,Mbf-1,Mep-1,Mtf-1
FT FT      /note= "Transcription factor binding site"
FT FT      2926..2933
FT FT      /*tag= ah
FT FT      /label= c/ebp
FT FT      /note= "Transcription factor binding site"
FT FT      2951..2956
FT FT      /*tag= ai
FT FT      /label= Pea-3
FT FT      /note= "Transcription factor binding site"
FT FT      2988..2993
FT FT      /*tag= aj
FT FT      /label= Myb
FT FT      /note= "Transcription factor binding site"
FT FT      2998..3003
FT FT      /*tag= ak
FT FT      /label= Pea-3
FT FT      /note= "Transcription factor binding site"
FT FT      3076..3086
FT FT      /*tag= al
FT FT      /label= SRE
FT FT      /note= "Transcription factor binding site"
FT FT      3102..3109
FT FT      /*tag= am
FT FT      /label= Ap-2
```